Secure Degrees of Freedom Region of the Two-User MISO Broadcast Channel with Alternating CSIT*

Pritam Mukherjee[1], Ravi Tandon[2], and Sennur Ulukus[1]

[1]Department of ECE, University of Maryland, College Park, MD
[2]Discovery Analytics Center and Department of CS, Virginia Tech, VA

February 10, 2015

Abstract

The two user multiple-input single-output (MISO) broadcast channel with confidential messages (BCCM) is studied in which the nature of channel state information at the transmitter (CSIT) from each user can be of the form I_i, $i = 1, 2$ where $I_1, I_2 \in \{\mathsf{P}, \mathsf{D}, \mathsf{N}\}$, and the forms P, D and N correspond to perfect and instantaneous, completely delayed, and no CSIT, respectively. Thus, the overall CSIT can alternate between 9 possible states corresponding to all possible values of $I_1 I_2$, with each state occurring for $\lambda_{I_1 I_2}$ fraction of the total duration. The main contribution of this paper is to establish the secure degrees of freedom (s.d.o.f.) region of the MISO BCCM with alternating CSIT with the symmetry assumption, where $\lambda_{I_1 I_2} = \lambda_{I_2 I_1}$.

The main technical contributions include developing a) novel achievable schemes for MISO BCCM with alternating CSIT with security constraints which also highlight the synergistic benefits of inter-state coding for secrecy, b) new converse proofs via local statistical equivalence and channel enhancement; and c) showing the interplay between various aspects of channel knowledge and their impact on s.d.o.f.

1 Introduction

Wireless systems are particularly vulnerable to security attacks because of the inherent openness of the transmission medium. With the widespread adoption of multiple-input multiple-output (MIMO) systems, there has been a significant recent interest in information theoretic

*This work was supported by NSF Grants CNS 13-14733, CCF 14-22111, CCF 14-22129 and CCF 14-22090, and presented in part at IEEE ISIT 2014 and to be presented in part at IEEE ICC 2015.

physical layer security, the main premise of which is to exploit the difference in the wireless channels between different users. Information theoretic security has been investigated for a variety of channel models ranging from fading channels [1–4], MIMO wiretap channels [5–8], multiple access channels [9–13], multi-receiver wiretap channels [14–16], broadcast channels with confidential messages [17–19], wiretap channels with helpers [20, 21], interference channels with confidential messages [22–25], X-channels with confidential messages [26, 27], relay eavesdropper channels [28–32], etc.

The focus of this paper is on the secure degrees of freedom (s.d.o.f.) region of the fading two-user multiple-input single-output (MISO) broadcast channel with confidential messages (BCCM), in which the transmitter with two antennas has two confidential messages, one for each of the single antenna users (see Fig. 1). The secrecy capacity region of the MISO broadcast channel for the case of perfect and instantaneous CSI at all terminals (transmitter and the receivers) has been characterized in [18, 19]. Using these results, it follows that for the two-user MISO BCCM, the sum s.d.o.f. is 2 with perfect and instantaneous channel state information at the transmitter (CSIT). In practice, the assumption of perfect and instantaneous CSIT may be too optimistic as CSIT may be delayed, imprecise or may not even be available at all.

The impact of relaxing such assumptions on the d.o.f. (secure or otherwise) has been widely studied in the literature. With perfect CSIT (P), the sum d.o.f. for the two-user MISO broadcast channel is 2. With no CSIT (N) however, reference [33] showed that the sum d.o.f.[1] collapses to 1. With delayed[2] CSIT (D), it is shown in [34] that the sum d.o.f. for the two-user MISO BC increases to $\frac{4}{3}$. [34] also presents novel results for the more general setting of K-user MISO BC, for $K \geq 2$. With delayed CSI, [35] established the d.o.f. region for the two-user MIMO BC. Other channel models besides the BC has also been investigated. Reference [36] provided the d.o.f. region of the MIMO interference channel with delayed CSIT and output feedback. For the X-channel, references [37, 38] showed that the optimal sum d.o.f. is $\frac{4}{3}$ with perfect channel knowledge. With delayed CSIT the optimal sum d.o.f. of the X-channel remains unknown in general. However, with a restriction of the transmission policies to linear schemes, reference [39] determined the sum d.o.f. of the channel to be $\frac{6}{5}$ (also see [40, 41] and the references therein). With global feedback, where each transmitter receives output feedback from every receiver, [42] showed that the sum d.o.f. of the two-user X-channel with delayed CSIT is the same as that of the two-user MISO broadcast channel with 2 antennas at the transmitter; thus, all the transmitters can cooperate and behave like a single 2-antenna MISO system, and the optimal sum d.o.f. is $\frac{4}{3}$.

[1] We refer to sum d.o.f. as the sum degrees of freedom for a network without any confidentiality constraints (e.g., MISO BC); and sum s.d.o.f. as the sum secure degrees of freedom for the same network with confidential messages (e.g., MISO BCCM).

[2] By delayed CSIT, we refer to the standard assumption as in [34] in which the delay in acquiring CSIT is larger than the channel coherence time.

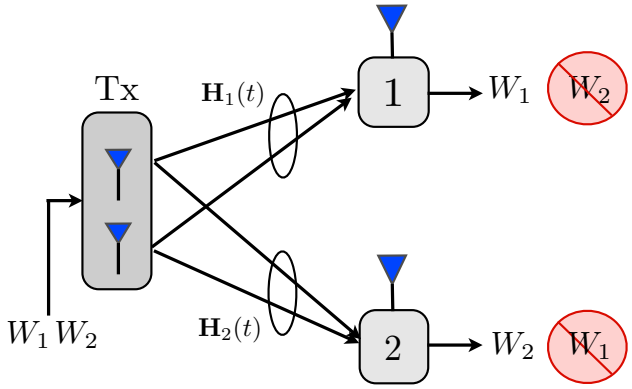


Figure 1: MISO broadcast channel with confidential messages (BCCM).

When security constraints are introduced, the s.d.o.f. is known for several scenarios of delayed or no CSIT. For the two-user MISO BCCM with no CSIT, the sum s.d.o.f. is zero as the two users are statistically equivalent and hence no secrecy is possible. On the other hand, with completely outdated CSIT from both users, [43] showed that the sum s.d.o.f. increases to 1. For the two-user SISO X-channel with confidential messages and global output feedback, [44] showed that the optimal s.d.o.f. is 1; thus, two distributed transmitters with one antenna each are as good as a single transmitter with 2 antennas and the X-channel behaves like a two-user MISO BCCM. The aforementioned literature primarily deals with homogeneous CSIT scenarios in which the nature of channel knowledge supplied by every receiver is of the same form. In practice, however, the nature of CSIT can vary across users. This observation naturally leads to the setting of heterogeneous (or hybrid) CSIT which models the variability in the quality/delay of channel knowledge supplied by different users. In contrast to homogeneous CSIT, the setting of heterogeneous CSIT is much less understood. To the best of our knowledge, the complete characterization of the d.o.f. of all fixed heterogeneous CSIT configurations is only known for the two-user MISO broadcast channel: see [45, 46] for state PD for which the optimal sum d.o.f. is shown to be 3/2; and [47] which recently settled the states PN and DN through a novel converse proof and showed that the optimal sum d.o.f. is given by 1. Beyond these results, partial results are available for the three-user MISO BC with hybrid CSIT in [48, 49] but by and large the problem of heterogeneous CSIT even without secrecy constraints remains open.

Besides exhibiting heterogeneity across users, the nature of channel knowledge may also vary over time/frequency. Such variability can arise either naturally (due to the time variation in tolerable feedback overhead from a user) or it can be artificially induced (by deliberately altering the channel feedback mechanism over time/frequency). For example, instead of requiring perfect CSIT from one user and delayed CSIT from the other user throughout the duration of communication, one may require that for half of the time, the first user provide perfect CSIT while the second user provide delayed CSIT (state PD), and the roles of the

users are reversed for the remaining half of the time (state DP), the total network feedback overhead being the same in both cases. This leads naturally to the setting of alternating CSIT in which multiple CSIT states, for instance, PD and DP in the above example, arise over time. The alternating CSIT framework was introduced in [50] where the d.o.f. region was characterized for the two-user MISO BC. It was shown that synergistic gains in d.o.f. are possible by jointly coding across these states. It was observed in [50] for the two-user case that the final d.o.f. region depends only on the marginal fractions of perfect, delayed and no CSIT, that is, the fractions of the time a user provides perfect, delayed and no CSIT. Given these results, several natural questions arise: a) do such synergistic gains still exist with additional confidentiality constraints on the messages, b) if yes, what is the optimal s.d.o.f. region and how to achieve it, c) what is the penalty for incorporating confidentiality in contrast to [50] and d) the fundamental impact of the variability of channel knowledge on secrecy.

In this paper, we consider the two-user MISO BCCM with alternating CSIT with all 9 possible CSIT states: PP, PD, PN, DP, NP, DD, DN, ND, and NN. We assume that these states occur for arbitrary fractions of time, except for a mild condition of symmetry, which is that states I_1I_2 and I_2I_1 occur for equal fractions of the time if $I_1 \neq I_2$. The main contribution of this paper is the characterization of the optimal s.d.o.f. region for this general model[3]. With 9 states, each occurring for arbitrary fractions of the time, it is not immediately clear how to optimally code across the states and the achievability of the s.d.o.f. region is highly non-trivial. To this end, we first develop several key constituent schemes, where each scheme uses a subset of the 9 states to achieve a particular s.d.o.f. value. We present all the constituent schemes in Section 4. Now given an arbitrary[4] probability mass function (pmf) on the 9 CSIT states, we need to judiciously time share between the constituent schemes to achieve the optimal s.d.o.f. region. It is not immediately clear how this should be done. Thus, we consider different sub-cases based on the relative proportions of the various states and explicitly characterize how the constituent schemes should be time shared to obtain the optimal s.d.o.f. region in each sub-case. This characterization is done in Section 5.

Next, we provide a matching converse for the full region. We first generalize the *local statistical equivalence* property introduced in [51]. The idea behind the converse is to first enhance the channel by providing more CSIT to obtain a new channel with fewer number of states but at least as large secrecy capacity as the original channel. Outer bounds on the s.d.o.f. region for the enhanced channel give us the desired outer bounds for the original channel.

Thus, the main contributions of this paper can be summarized as follows: a) We obtain

[3]In our preliminary work [51], we considered the problem with only two states, PD and DP and established the optimal s.d.o.f. region for this specific problem. Reference [52] considered another special case with four states: PP, PD, DP and DD, but provided only an inner bound for the s.d.o.f. region.

[4]Arbitrary subject to mild symmetry, i.e., $\lambda_{I_1I_2} = \lambda_{I_2I_1}$

the full s.d.o.f. region with all possible 9 states occurring for arbitrary fractions of time constrained only by the requirement of symmetry, which is that states I_1I_2 and I_2I_1 occur for equal fractions of the time if $I_1 \neq I_2$. b) To achieve this region, we provide several new optimal achievable schemes for different alternating CSIT scenarios. c) In addition, we provide an explicit method of combining the various achievable schemes judiciously to achieve the region. d) We provide a matching converse for the full region using channel enhancement and generalizing the *local statistical equivalence* property introduced in [51]. e) We establish the s.d.o.f. regions of the MISO BCCM under two heterogeneous CSIT settings: PD and DN states alone. These results completely settle the problem of characterizing the s.d.o.f. regions of all individual heterogenous CSIT states: PD, PN, DN. f) We show synergistic benefits of coding across the different alternating states even under security constraints.

2 System Model

We consider a two-user MISO BC, shown in Fig. 1, where the transmitter Tx, equipped with 2 antennas, wishes to send independent confidential messages to two single antenna receivers 1 and 2. The input-output relations at time t are given by,

$$Y(t) = \mathbf{H}_1(t)\mathbf{X}(t) + N_1(t) \tag{1}$$

$$Z(t) = \mathbf{H}_2(t)\mathbf{X}(t) + N_2(t), \tag{2}$$

where $Y(t)$ and $Z(t)$ are the channel outputs of receivers 1 and 2, respectively. The 2×1 channel input $\mathbf{X}(t)$ is power constrained as $\mathbb{E}[||\mathbf{X}(t)||^2] \leq P$, and $N_1(t)$ and $N_2(t)$ are circularly symmetric complex white Gaussian noises with zero-mean and unit-variance. The 1×2 channel vectors $\mathbf{H}_1(t)$ and $\mathbf{H}_2(t)$ of receivers 1 and 2, respectively, are independent and identically distributed (i.i.d.) with continuous distributions, and are also i.i.d. over time. We denote $\mathbf{H}(t) = \{\mathbf{H}_1(t), \mathbf{H}_2(t)\}$ as the collective channel vectors at time t and $\mathbf{H}^n = \{\mathbf{H}(1), \ldots, \mathbf{H}(n)\}$ as the sequence of channel vectors up until and including time n.

In practice, the receivers estimate the channel coefficients and feed them back to the transmitter. In general, the receiver can choose to send not only the current measurements, but rather any function of all the channel measurements it has taken upto that time. The CSIT at time t can thus be any function of the measured channel coefficients upto time t. There are two key aspects to the CSIT: precision and delay. Precision captures the fact that the measurements made at the receivers and sent to the transmitter are imprecise (usually, quantized) and noisy. Delay is introduced since making measurements and feeding them back to the transmitter takes time. We will focus on the delay aspect of CSIT, and assume that the CSIT when available, has infinite precision.

In order to model the delay in CSIT, we assume that at each time t, there are 3 possible

CSIT states for each user:

- *Perfect CSIT* (P): This denotes the availability of precise and instantaneous CSI of a user at the transmitter. In this state, the transmitter has precise channel knowledge before the start of the communication.
- *Delayed CSIT* (D): In this state, the transmitter does not have the CSI at the beginning of the communication. In slot t, the receiver may send any function of all the channel coefficients upto and including time t as CSI to the transmitter. However, the CSIT becomes available only after a delay such that the CSI is completely outdated, that is, independent of the current channel realization.
- *No CSIT* (N): In this state, there is no CSI of the user available at the transmitter.

Denote the CSIT of user 1 by I_1 and the CSIT of user 2 by I_2. Then,

$$I_1, I_2 \in \{\mathsf{P}, \mathsf{D}, \mathsf{N}\}. \tag{3}$$

Thus, for the two-user MISO BC, we have 9 CSIT states, namely PP, DD, NN, PD, DP, PN, NP, DN, and ND. Let $\lambda_{I_1 I_2}$ be the fraction of the time the state $I_1 I_2$ occurs. Then,

$$\sum_{I_1, I_2} \lambda_{I_1 I_2} = 1. \tag{4}$$

We also assume symmetry: $\lambda_{I_1 I_2} = \lambda_{I_2 I_1}$ for every $I_1 I_2$. Specifically,

$$\lambda_{PD} = \lambda_{DP} \tag{5}$$

$$\lambda_{DN} = \lambda_{ND} \tag{6}$$

$$\lambda_{PN} = \lambda_{NP}. \tag{7}$$

Further, we assume that perfect and global CSI is available at both receivers.

A secure rate pair (R_1, R_2) is achievable if there exists a sequence of codes which satisfy the reliability constraints at the receivers, namely, $\Pr\left[W_i \neq \hat{W}_i\right] \leq \epsilon_n$, for $i = 1, 2$, and the confidentiality constraints, namely,

$$\frac{1}{n} I(W_1; Z^n, \mathbf{H}^n) \leq \epsilon_n, \qquad \frac{1}{n} I(W_2; Y^n, \mathbf{H}^n) \leq \epsilon_n, \tag{8}$$

where $\epsilon_n \to 0$ as $n \to \infty$. Informally, the constraints in (8) ensure that the information leakage, per channel use, of the first receiver's message at the second receiver should be arbitrarily small, and vice versa. A s.d.o.f. pair (d_1, d_2) is achievable, if there exists an

achievable rate pair (R_1, R_2) such that

$$d_1 = \lim_{P\to\infty} \frac{R_1}{\log P}, \qquad d_2 = \lim_{P\to\infty} \frac{R_2}{\log P}. \tag{9}$$

Let us define the following:

$$\lambda_P \triangleq \lambda_{PP} + \lambda_{PD} + \lambda_{PN} \tag{10}$$
$$\lambda_D \triangleq \lambda_{PD} + \lambda_{DD} + \lambda_{DN} \tag{11}$$
$$\lambda_N \triangleq \lambda_{PN} + \lambda_{DN} + \lambda_{NN}. \tag{12}$$

Using these definitions, it is easy to verify that

$$\lambda_P + \lambda_D + \lambda_N = 1. \tag{13}$$

Here, we can interpret these three quantities as follows:

- λ_P: represents the total fraction of time the CSIT of a user is in the P state.
- λ_D: represents the total fraction of time the CSIT of a user is delayed, that is, the state D.
- λ_N: represents the total fraction of time a user supplies no CSIT.

Given the probability mass function (pmf), $\lambda_{I_1 I_2}$, our goal is to characterize the s.d.o.f. region of the two-user MISO BCCM.

3 Main Result and Discussion

Theorem 1 *The s.d.o.f. region for the two-user MISO BCCM with alternating CSIT, $\mathcal{D}(\lambda_{I_1 I_2})$, is the set of all non-negative pairs (d_1, d_2) satisfying,*

$$d_1 \le \min\left(\frac{2 + 2\lambda_P - \lambda_{PP}}{3}, 1 - \lambda_{NN}\right) \tag{14}$$
$$d_2 \le \min\left(\frac{2 + 2\lambda_P - \lambda_{PP}}{3}, 1 - \lambda_{NN}\right) \tag{15}$$
$$3d_1 + d_2 \le 2 + 2\lambda_P \tag{16}$$
$$d_1 + 3d_2 \le 2 + 2\lambda_P \tag{17}$$
$$d_1 + d_2 \le 2(\lambda_P + \lambda_D). \tag{18}$$

A proof for the achievability of this region will be provided in Section 5 using constituent schemes presented in Section 4. A converse is provided in Section 6.

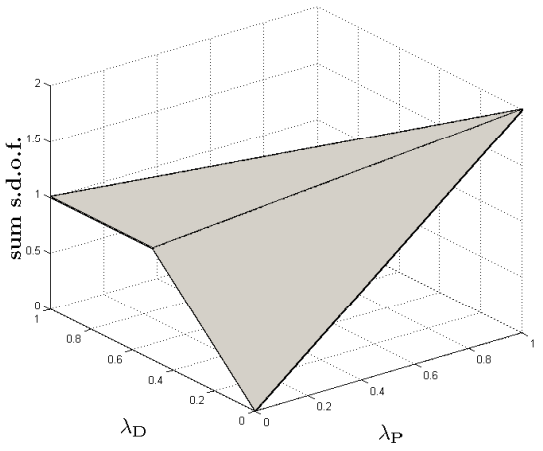


Figure 2: The sum s.d.o.f. as a function of λ_P and λ_D.

We next make a series of remarks highlighting the consequences and interesting aspects of this theorem.

Remark 1. [Sum s.d.o.f.: $\max(d_1 + d_2)$]

From the region stated in (14)-(18), it is clear that the sum s.d.o.f. is given by,

$$\text{sum s.d.o.f.} = \min\left(2\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}\right), 2(1-\lambda_{NN}), 2(\lambda_P+\lambda_D), 1+\lambda_P\right). \tag{19}$$

The sum s.d.o.f. expression in (19) can be significantly simplified by noting that the first two terms in the minimum are inactive due to the inequalities $1+\lambda_P \leq 2\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}\right)$, and $2(\lambda_P+\lambda_D) = 2(1-\lambda_N) \leq 2(1-\lambda_{NN})$. These inequalities follow directly from (10)-(13). Using these inequalities, the sum s.d.o.f. expression above is equivalent to

$$\begin{aligned}\text{sum s.d.o.f.} &= \min\left(2(\lambda_P+\lambda_D), 1+\lambda_P\right) &(20)\\ &= \min\left(2(\lambda_P+\lambda_D), 2\lambda_P+\lambda_D+\lambda_N\right) &(21)\\ &= 2\lambda_P+\lambda_D+\min(\lambda_D,\lambda_N). &(22)\end{aligned}$$

Fig. 2 shows the sum s.d.o.f. as a function of λ_P and λ_D.

Remark 2. [Same marginals property]

From (22), we notice that the marginal probabilities λ_P, λ_D and λ_N are sufficient to determine the sum s.d.o.f. *Thus, for any given pmf* $\lambda_{I_1I_2}$*, satisfying the symmetry conditions* (5)-(7)*, there exists an* ***equivalent*** *alternating CSIT problem having only three states:* PP*,* DD *and* NN *occurring for* λ_P*,* λ_D *and* λ_N *fractions of the time, respectively, that has the same sum*

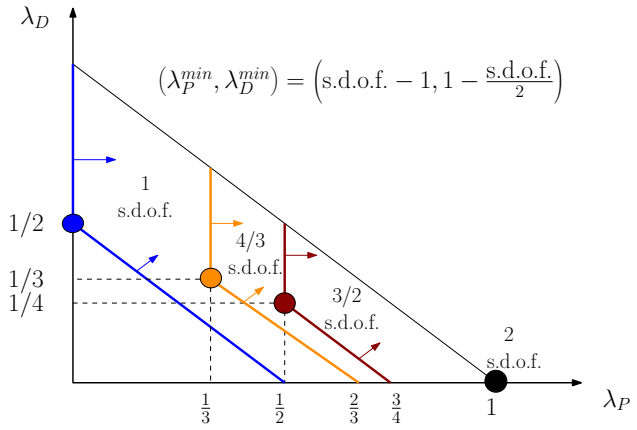


Figure 3: Trade-off between delayed and perfect CSIT.

s.d.o.f. This observation is similar to the case when there is no secrecy [50]. *However unlike in [50], the s.d.o.f. region* ***does not*** *have the same property in general* as we can see the explicit dependence of the s.d.o.f. region in (14)-(18) on λ_{PP} and λ_{NN}.

Remark 3. [Channel knowledge equivalence]

We next highlight an interesting property which shows that from the sum s.d.o.f. perspective, no CSIT is equivalent to delayed CSIT when $\lambda_D \geq \lambda_N$, and delayed CSIT is equivalent to perfect CSIT when $\lambda_D < \lambda_N$.

Equivalence of delayed and no CSIT when $\lambda_D \geq \lambda_N$: From a sum s.d.o.f. perspective, we see that when $\lambda_D \geq \lambda_N$, the sum s.d.o.f. depends only on λ_P. Hence, as long as $\lambda_D \geq \lambda_N$ holds, the N states behave as D states in the sense that, if the N states were enhanced to D states, the sum s.d.o.f. would not increase. Essentially, the N states can be combined with various D states and we obtain the same sum s.d.o.f. as if every N state were replaced by a D state. Consider an example, where the states PD, DP and NN occur for $\frac{2}{5}$th, $\frac{2}{5}$th and $\frac{1}{5}$th fractions of the time, respectively. Note that $\lambda_D = \frac{2}{5} > \lambda_N = \frac{1}{5}$ in this case. The sum s.d.o.f., from (22), is $2\lambda_P + \lambda_D + \lambda_N = \frac{7}{5}$. Now, if we enhance the N states to D states, we get the states PD, DP and DD occur for $\frac{2}{5}$th, $\frac{2}{5}$th and $\frac{1}{5}$th of the time, respectively. The sum s.d.o.f. of this enhanced system is still $\frac{7}{5}$.

Equivalence of delayed and perfect CSIT when $\lambda_D \leq \lambda_N$: From a sum s.d.o.f. perspective, we see that when $\lambda_D \leq \lambda_N$, the sum s.d.o.f. depends only on λ_N. Hence, in this case, if $\lambda_D \leq \lambda_N$, the delayed CSIT is as good as perfect CSIT, that is, every D state can be enhanced to a P state without any increase in the sum s.d.o.f. For example, consider a system where the states PD, DP and NN occur for $\frac{1}{5}$th, $\frac{1}{5}$th and $\frac{3}{5}$th fractions of the time, respectively. Note that $\lambda_D = \frac{1}{5} < \lambda_N = \frac{3}{5}$ in this case. The sum s.d.o.f. for this system is $\frac{4}{5}$, from (22). By enhancing the D states to P states, we get a system, where the states PP and NN occur for $\frac{2}{5}$th and $\frac{3}{5}$th fractions of the time, respectively. The sum s.d.o.f. in for this enhanced system

is still $\frac{4}{5}$.

Remark 4. [Minimum CSIT required for a sum s.d.o.f. value]

Fig. 3 shows the trade-off between λ_P and λ_D for a given value of sum s.d.o.f. The highlighted corner point in each curve shows the most *efficient* point in terms of CSIT requirement. *Any other feasible point either involves redundant CSIT or unnecessary instantaneous CSIT where delayed CSIT would have sufficed.* For example, following are the minimum CSIT requirements for various sum s.d.o.f. values:

$$\text{sum s.d.o.f.} = 2 \ : (\lambda_P, \lambda_D)_{\min} = (1, 0) \tag{23}$$

$$\text{sum s.d.o.f.} = \frac{3}{2} : (\lambda_P, \lambda_D)_{\min} = \left(\frac{1}{2}, \frac{1}{4}\right) \tag{24}$$

$$\text{sum s.d.o.f.} = \frac{4}{3} : (\lambda_P, \lambda_D)_{\min} = \left(\frac{1}{3}, \frac{1}{3}\right) \tag{25}$$

$$\text{sum s.d.o.f.} = 1 \ : (\lambda_P, \lambda_D)_{\min} = \left(0, \frac{1}{2}\right). \tag{26}$$

In general, for a given value of sum s.d.o.f. $= s$, the minimum CSIT requirements are given by:

$$(\lambda_P, \lambda_D)_{\min} = \begin{cases} \left(s-1, 1-\frac{s}{2}\right), & \text{if } 1 \le s \le 2 \\ \left(0, \frac{s}{2}\right), & \text{if } 0 \le s \le 1. \end{cases} \tag{27}$$

Remark 5. [Cost of security]

We recall that in the case with no security [50], the sum d.o.f. is given by,

$$\text{sum d.o.f.} = 2 - \frac{2\lambda_N}{3} - \frac{\max(\lambda_N, 2\lambda_D)}{3}. \tag{28}$$

Comparing with (22), we see that the loss in d.o.f. that must be incurred to incorporate secrecy constraints is given by,

$$(\text{sum d.o.f.}) - (\text{sum s.d.o.f.}) \triangleq \text{loss} = \begin{cases} \lambda_N, & \text{if } \lambda_N \ge 2\lambda_D \\ \frac{2}{3}(2\lambda_N - \lambda_D), & \text{if } 2\lambda_D \ge \lambda_N \ge \lambda_D \\ \frac{1}{3}(\lambda_N + \lambda_D), & \text{if } \lambda_D \ge \lambda_N. \end{cases} \tag{29}$$

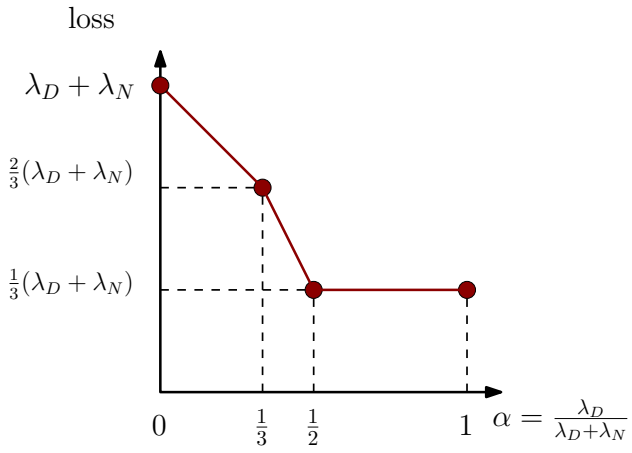


Figure 4: Cost of security.

If we define $\alpha = \lambda_D/(\lambda_D + \lambda_N)$, we can rewrite (29) as follows,

$$\text{loss} = (\lambda_D + \lambda_N) \times \begin{cases} (1-\alpha), & \text{if } \alpha \leq \frac{1}{3} \\ \left(\frac{4}{3} - 2\alpha\right), & \text{if } \frac{1}{2} \geq \alpha \geq \frac{1}{3} \\ \frac{1}{3}, & \text{if } \alpha \geq \frac{1}{2}. \end{cases} \tag{30}$$

We show this loss as a function of α in Fig. 4. Note that $\lambda_D + \lambda_N$ is the fraction of the time a user feeds back imperfect (delayed or none) CSIT. If this fraction is fixed, increasing the fraction of the delayed CSIT decreases the penalty due to the security constraints, but only to a certain extent. When $\lambda_N \geq \lambda_D$, increasing the fraction of delayed CSIT leads to a decrease in the penalty due to the security constraints. However, once the fraction of the delayed CSIT (state D) matches that of no CSIT (N), that is, $\lambda_D \geq \lambda_N$, increasing the fraction of delayed CSIT further does not reduce the penalty any more.

Remark 6. [S.d.o.f. characterization of individual CSIT states]

As an additional relevant result, we also characterize the respective s.d.o.f. regions for the 6 individual CSIT states. To the best of our knowledge, the only CSIT states for which the s.d.o.f. regions were previously known are: PP (with sum s.d.o.f.= 2), DD (with sum s.d.o.f.= 1), PN (with s.d.o.f.= 1), and NN (with s.d.o.f.= 0). For the remaining two CSIT states, i.e., PD and DN, we establish the optimal s.d.o.f. regions. In particular, for the PD CSIT state, we show in Appendix D that the s.d.o.f. region is given by $d_1 + d_2 \leq 1$. For the DN state, we show in Appendix E that the s.d.o.f. region is given by $d_1 + d_2 \leq 1/2$. As the next remark shows, these complete set of results for the individual CSIT states confirm the synergistic benefits (or lack thereof) in various alternating CSIT scenarios.

Remark 7. [Synergistic benefits]

It was shown in [50] that by coding across different states one can achieve higher sum d.o.f. than by optimal encoding for each state separately and time sharing. A similar result holds true in our case as well. We illustrate this with the help of a few examples.

Example 1. Consider a special case where only states PD and DP occur, each for half of the time. In our previous work, [51], we showed that optimal sum s.d.o.f. is $\frac{3}{2}$ in this case; see also (22) here. The best achievable scheme for the PD (or DP) state alone was known to achieve a sum s.d.o.f. of 1. This was either by treating the PD state as a PN state and zero forcing, or by treating PD as a DD state. However a converse proof showing the optimality of 1 sum s.d.o.f. was not known. In Appendix D, we present a converse proof to show that the sum s.d.o.f. of 1 is indeed optimal for the PD state alone. Thus, by encoding for each state separately and time sharing between the PD and DP states, we can achieve only 1 sum s.d.o.f., whereas joint encoding across the states achieves sum s.d.o.f. of $\frac{3}{2}$. Thus, we have synergistic benefit of 50% in this case.

Example 2. Consider another special case with three states: PD, DP and NN each occurring for one-third of the time. The optimal sum s.d.o.f. is $\frac{4}{3}$. If we encode for each state separately and time share between them, we can achieve a sum s.d.o.f. of $\frac{1}{3}\times 1+\frac{1}{3}\times 1+\frac{1}{3}\times 0 = \frac{2}{3}$, since the NN state does not provide any secrecy. If we encode across the PD and DP states optimally and then time share with the NN state, we can achieve $\frac{2}{3}\times\frac{3}{2}+\frac{1}{3}\times 0 = 1$ sum s.d.o.f. Thus, in this case too, we get synergistic benefit by coding across all the states together.

Example 3. Now, assume we have the following three states: PN, NP and DD each occurring for one-third of the time. The optimal sum s.d.o.f. for this case is $\frac{4}{3}$. On the other hand, the optimal sum s.d.o.f. of the PN state alone is 1, [47], and that of the DD state alone is also 1, [43]. Thus, by separately encoding for each state and time sharing, we can achieve $\frac{1}{3}\times 1+\frac{1}{3}\times 1+\frac{1}{3}\times 1 = 1$ sum s.d.o.f. Note that the optimal sum s.d.o.f. for PN and NP states, each occurring for half of the time, is also 1, using (22). Thus, by optimal encoding for PN and NP together and time sharing with the DD state also yields sum s.d.o.f. of 1. Therefore, there is synergistic benefit to be gained by coding across all the states together in this case too.

Example 4. Consider the case where the two states, DD and NN occur for equal fractions of time. The optimal sum s.d.o.f. of the DD state alone is 1 [43]. The NN state, by itself does not provide any secrecy and its s.d.o.f. $= 0$. Thus, by encoding for the individual states and time sharing, at most $1\times\frac{1}{2}+0\times\frac{1}{2} = \frac{1}{2}$ sum s.d.o.f. is achievable. However, by jointly encoding across both the DD and NN states, the optimal sum s.d.o.f. of 1 is achievable. Thus, we have synergistic benefit of 100% in terms of sum s.d.o.f. in this case.

Example 5. Finally, consider the case where the two states, DN and ND occur for equal fractions of time. We show in Appendix E that the optimal sum s.d.o.f. for DN state is $\frac{1}{2}$. Thus, by separately encoding across the individual states, only $\frac{1}{2}$ sum s.d.o.f. is achievable.

However, by jointly encoding across both the DN and DN states, the optimal sum s.d.o.f. of 1 is achievable. Thus, we have synergistic benefit of 100% in terms of sum s.d.o.f. in this case.

Remark 7. [Lack of synergistic benefits]

There are some situations where joint encoding across alternating states does not yield any benefit in terms of the s.d.o.f. region. For example, consider a case with only 2 states, PN and NP, each occurring for half of the time. The optimal sum s.d.o.f. for the PN state alone is 1, which is achieved by zero forcing. The optimal sum s.d.o.f. of both PN and NP states together is also 1; thus, encoding for each state separately is optimal in this case. Indeed separable encoding for each individual state suffices to achieve the full s.d.o.f. region as well. *This result is perhaps surprising, since in the case with no security, we do get synergistic benefits of joint encoding across the* PN *and* NP *states. The optimal sum s.d.o.f. with joint encoding is* $\frac{3}{2}$*, while that for each state alone is* 1*,* [50].

4 Constituent Schemes

Summary of Constituent Schemes (CS)				
Sum s.d.o.f.	CS Notation	CSIT States	Fractions of States	(d_1, d_2)
2	S^2	PP	1	$(1,1)$
3/2	$S_1^{3/2}$	PD, DP	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\left(\frac{3}{4}, \frac{3}{4}\right)$
	$S_2^{3/2}$	PD, DP, PN, NP	$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$	$\left(\frac{3}{4}, \frac{3}{4}\right)$
4/3	$S_1^{4/3}$	PD, DP, NN	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	$\left(\frac{2}{3}, \frac{2}{3}\right)$
	$S_2^{4/3}$	PN, NP, DD	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	$\left(\frac{2}{3}, \frac{2}{3}\right)$
1	S_1^1	DD	1	$\left(\frac{1}{2}, \frac{1}{2}\right)$
	S_2^1	DD, NN	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$
	S_3^1	DN, ND	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$
2/3	$S_1^{2/3}$	DD	1	$\left(\frac{2}{3}, 0\right)$
	$S_2^{2/3}$	DD, NN	$\left(\frac{2}{3}, \frac{1}{3}\right)$	$\left(\frac{2}{3}, 0\right)$
	$S_3^{2/3}$	DN, ND, NN	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	$\left(\frac{2}{3}, 0\right)$

Table 1: Constituent schemes.

Before we present the achievability of the s.d.o.f. region, we first present the key constituent schemes that will be instrumental in the proof. We combine these schemes carefully

and time share between them to achieve the s.d.o.f. region. A summary of these constituent schemes is shown in Table 1. Before we discuss the individual schemes we make the following remark that applies to all the schemes presented here.

4.1 A Note on the Achievable Security Guarantee

Each scheme described in the following sections can be outlined as follows. We neglect the impact of noise at high SNR. Then, to achieve a certain s.d.o.f. pair (d_1, d_2), we send n_1 symbols $\underline{\mathbf{u}} = (u_1, \ldots, u_{n_1})$ and n_2 symbols $\underline{\mathbf{v}} = (v_1, \ldots, v_{n_2})$ intended for the first and second receivers, respectively, in n_B slots, such that $d_1 = n_1/n_B$ and $d_2 = n_2/n_B$. Finally, we argue that the leakage of information symbols at the unintended receiver is $o(\log P)$. We however want a stronger guarantee of security, namely,

$$\frac{1}{n} I(W_1; Z^n, \mathbf{H}^n) \leq \epsilon_n, \qquad \frac{1}{n} I(W_2; Y^n, \mathbf{H}^n) \leq \epsilon_n. \tag{31}$$

To achieve this, we view the n_B slots described in the scheme as a block and treat the equivalent channel from $\underline{\mathbf{u}}$ to $(\mathbf{Y}, \mathbf{H})$ and $(\mathbf{Z}, \mathbf{H})$ as a memoryless wiretap channel (with $(\mathbf{Y}, \mathbf{H})$ being the legitimate receiver) by ignoring the CSI of the previous block. We do the same for the channel from $\underline{\mathbf{v}}$ to $(\mathbf{Z}, \mathbf{H})$ and $(\mathbf{Y}, \mathbf{H})$ (with $(\mathbf{Z}, \mathbf{H})$ as the legitimate receiver). Note also that no information about $\mathbf{H}$ is used to create the codebooks for $\underline{\mathbf{u}}$ and $\underline{\mathbf{v}}$ in any of the schemes. More formally, the following secrecy rate pair is achievable for receivers 1 and 2, respectively, from [53]:

$$R_1 = I(\underline{\mathbf{u}}; \mathbf{Y}, \mathbf{H}) - I(\underline{\mathbf{v}}; \mathbf{Z}, \mathbf{H}) = I(\underline{\mathbf{u}}; \mathbf{Y}|\mathbf{H}) - I(\underline{\mathbf{v}}; \mathbf{Z}|\mathbf{H}) \tag{32}$$

$$R_2 = I(\underline{\mathbf{v}}; \mathbf{Z}, \mathbf{H}) - I(\underline{\mathbf{u}}; \mathbf{Y}, \mathbf{H}) = I(\underline{\mathbf{v}}; \mathbf{Z}|\mathbf{H}) - I(\underline{\mathbf{u}}; \mathbf{Y}|\mathbf{H}), \tag{33}$$

where we noted that $\underline{\mathbf{u}}$ and $\underline{\mathbf{v}}$ are all independent of $\mathbf{H}$. Using the proposed scheme, $\underline{\mathbf{u}}$ (resp., $\underline{\mathbf{v}}$) can be reconstructed from $(\mathbf{Y}, \mathbf{H})$ (resp., $(\mathbf{Z}, \mathbf{H})$) to within a noise distortion. Thus,

$$I(\underline{\mathbf{u}}; \mathbf{Y}|\mathbf{H}) = n_1 \log P + o(\log P) \tag{34}$$

$$I(\underline{\mathbf{v}}; \mathbf{Z}|\mathbf{H}) = n_2 \log P + o(\log P). \tag{35}$$

Also, for each scheme,

$$I(\underline{\mathbf{v}}; \mathbf{Y}|\mathbf{H}) = o(\log P) \tag{36}$$

$$I(\underline{\mathbf{u}}; \mathbf{Z}|\mathbf{H}) = o(\log P). \tag{37}$$

Thus, from (32) and (33), the achievable secure rates in each block are,

$$R_1 = n_1 \log P + o(\log P) \tag{38}$$

$$R_2 = n_2 \log P + o(\log P). \tag{39}$$

Since our block contains n_B channel uses, the effective secure rates are

$$R_1 = \frac{n_1}{n_B} \log P + o(\log P) \tag{40}$$

$$R_2 = \frac{n_2}{n_B} \log P + o(\log P). \tag{41}$$

These rates clearly yield the required s.d.o.f. pair (d_1, d_2), while also conforming to our stringent security requirement.

In the following subsections, we now present the achievability of each scheme in detail.

Notation: A particular sum s.d.o.f. value can be achieved in various ways through alternation between different possible sets of CSIT states. To this end, we use the following notation: if there are r schemes achieving a particular s.d.o.f. value, we denote these schemes as: $S_1^{\text{sum s.d.o.f.}}, S_2^{\text{sum s.d.o.f.}}, \ldots, S_r^{\text{sum s.d.o.f.}}$. For example, in Table 1, for achieving the sum s.d.o.f. value of 1, we present $r = 3$ distinct schemes and these are denoted as S_1^1, S_2^1 and S_3^1.

Given a 1×2 channel vector $\mathbf{H}(t)$, we denote by $\mathbf{H}(t)^\perp$, a 2×1 beamforming vector that is orthogonal to the 1×2 channel vector $\mathbf{H}(t)$; in other words, $\mathbf{H}(t)\mathbf{H}(t)^\perp = 0$.

4.2 Scheme Achieving Sum s.d.o.f. of 2

A sum s.d.o.f. of 2 is achievable only in the state PP, that is, when the transmitter has perfect CSIT from both users. This is achievable using zero-forcing. The following scheme achieves a sum s.d.o.f. of 2.

4.2.1 Scheme S^2

The scheme S^2 uses the state PP and achieves the rate pair $(d_1, d_2) = (1, 1)$. The scheme is as follows. We wish to send confidential symbols u and v to receivers 1 and 2, respectively, in one time slot, thus achieving a sum s.d.o.f. of 2. Since the transmitter knows both channel coefficients $\mathbf{H}_1$ and $\mathbf{H}_2$, it sends,

$$\mathbf{X} = u\mathbf{H}_2^\perp + v\mathbf{H}_1^\perp, \tag{42}$$

where, $\mathbf{H}_i(t)^\perp$ is a 2×1 beamforming vector that is orthogonal to the 1×2 channel vector $\mathbf{H}_i(t)$ for $i = 1, 2$. This is to ensure that the symbols do not leak to unintended receivers.

For s.d.o.f. calculations, we disregard the additive noise and the outputs at the receivers are:

$$Y = u\mathbf{H}_1\mathbf{H}_2^{\perp} \tag{43}$$

$$Z = v\mathbf{H}_2\mathbf{H}_1^{\perp}, \tag{44}$$

which allows both receivers to decode their respective messages. Also, since u does not appear at all in Z, the confidentiality of u is guaranteed. Similarly, the confidentiality of v too is satisfied.

4.3 Schemes Achieving Sum s.d.o.f. of 3/2

The following schemes achieve $\frac{3}{2}$ sum s.d.o.f.:

4.3.1 Scheme $S_1^{3/2}$

In this subsection, we present the scheme $S_1^{3/2}$ which uses the states $(\mathsf{PD}, \mathsf{DP})$ with fractions $(\frac{1}{2}, \frac{1}{2})$ to achieve rate pair $(d_1, d_2) = (\frac{3}{4}, \frac{3}{4})$.

This scheme was presented in [51]. For the sake of completeness we reproduce the scheme here. We wish to send 3 confidential symbols from the transmitter to each of the receivers in 4 channel uses at high P (that is negligible noise). Let us denote by (u_1, u_2, u_3) and (v_1, v_2, v_3) the confidential symbols intended for receivers 1 and 2, respectively. Also, in 2 of the 4 channel uses, the channel is in state PD; in the remaining 2 uses, the channel is in state DP. The scheme is as follows:

1) At time $t = 1$, $S(1) = \mathsf{PD}$: As the transmitter knows $\mathbf{H}_1(1)$, it sends:

$$\mathbf{X}(1) = [u_1 \quad 0]^T + q\mathbf{H}_1(1)^{\perp}, \tag{45}$$

where $\mathbf{H}_1(1)\mathbf{H}_1(1)^{\perp} = 0$, and q denotes an artificial noise distributed as $\mathcal{CN}(0, P)$. Here $\mathbf{H}_1(1)^{\perp}$ is a 2×1 beamforming vector orthogonal to the 1×2 channel vector $\mathbf{H}_1(1)$ of receiver 1 that ensures that the artificial noise q does not create interference at receiver 1. The receivers' outputs are:

$$Y(1) = h_{11}(1)u_1 \tag{46}$$

$$Z(1) = h_{21}(1)u_1 + q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp} \triangleq K. \tag{47}$$

Thus, receiver 1 has observed u_1 while receiver 2 gets a linear combination of u_1 and q, which we denote as K. Due to delayed CSIT from receiver 2, the transmitter can reconstruct K in the next channel use and use it for transmission.

2) At time $t = 2$, $S(2) = \mathsf{DP}$: The transmitter knows $\mathbf{H}_2(2)$ and K. It sends

$$\mathbf{X}(2) = [v_1 + K \quad v_2 + K]^T + u_2\mathbf{H}_2(2)^{\perp}. \tag{48}$$

The received signals are:

$$\begin{aligned} Y(2) &= h_{11}(2)v_1 + h_{12}(2)v_2 + (h_{11}(2) + h_{12}(2))K + u_2\mathbf{H}_1(2)\mathbf{H}_2(2)^{\perp} && (49)\\ &= L_1(v_1, v_2, K) + u_2\mathbf{H}_1(2)\mathbf{H}_2(2)^{\perp} && (50)\\ Z(2) &= h_{21}(2)v_1 + h_{22}(2)v_2 + (h_{21}(2) + h_{22}(2))K && (51)\\ &\stackrel{\Delta}{=} L_2(v_1, v_2, K), && (52) \end{aligned}$$

where we have defined $L_1(v_1, v_2, K)$ and $L_2(v_1, v_2, K)$ as linear combinations of v_1, v_2 and K at receivers 1 and 2, respectively.

3) At time $t = 3$, $S(3) = \mathsf{DP}$: The transmitter knows $\mathbf{H}_2(3)$ and $L_1(v_1, v_2, K)$ (via delayed CSIT from $t = 2$). Using these, it transmits:

$$\mathbf{X}(3) = [L_1(v_1, v_2, K) \quad 0]^T + u_3\mathbf{H}_2(3)^{\perp}, \tag{53}$$

and the channel outputs are:

$$\begin{aligned} Y(3) &= h_{11}(3)L_1(v_1, v_2, K) + u_3\mathbf{H}_1(3)\mathbf{H}_2(3)^{\perp} && (54)\\ Z(3) &= h_{21}(3)L_1(v_1, v_2, K). && (55) \end{aligned}$$

At the end of this step, note that, receiver 2 can decode v_1 and v_2 by first eliminating K using $Z(1)$ and $Z(3)$ to get a linear combination of v_1 and v_2, which it can then use with $Z(2)$ to solve for v_1 and v_2.

4) At time $t = 4$, $S(4) = \mathsf{PD}$: The transmitter knows $\mathbf{H}_1(4)$ and it sends

$$\mathbf{X}(4) = [L_1(v_1, v_2, K) \quad 0]^T + v_3\mathbf{H}_1(4)^{\perp}, \tag{56}$$

and the channel outputs are:

$$\begin{aligned} Y(4) &= h_{11}(4)L_1(v_1, v_2, K) && (57)\\ Z(4) &= h_{21}(4)L_1(v_1, v_2, K) + v_3\mathbf{H}_2(4)\mathbf{H}_1(4)^{\perp}. && (58) \end{aligned}$$

Thus, at the end of these four steps the outputs at the two receivers can be summarized

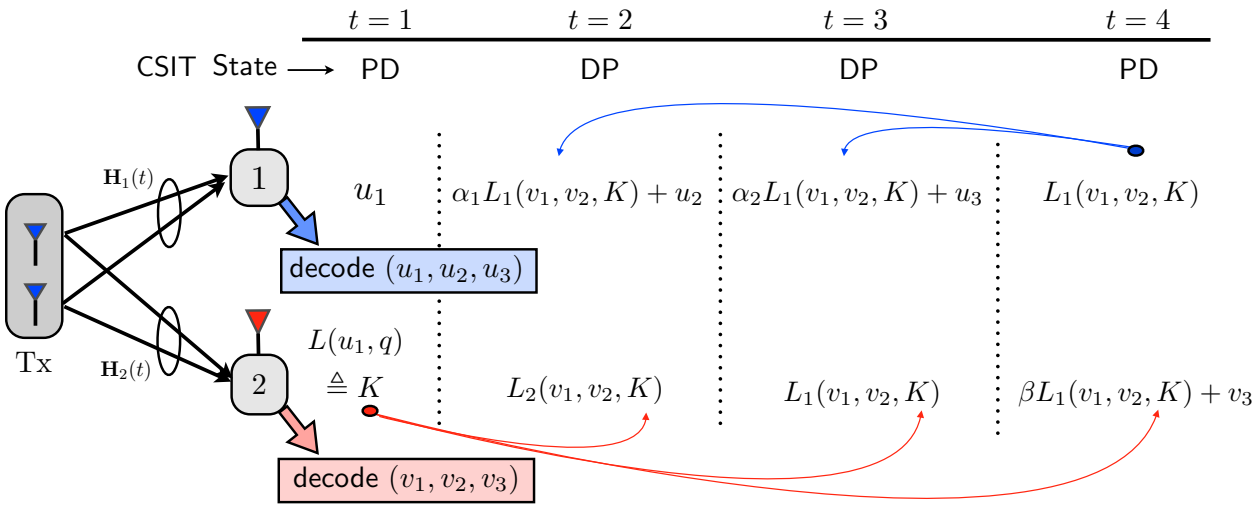


Figure 5: Achieving $\frac{3}{2}$ s.d.o.f. using scheme $S_1^{3/2}$.

(see Fig. 5) as:

$$\mathbf{Y}=\begin{bmatrix} u_1 \\ \alpha_1 L_1(v_1,v_2,K)+u_2 \\ \alpha_2 L_1(v_1,v_2,K)+u_3 \\ L_1(v_1,v_2,K) \end{bmatrix}, \qquad \mathbf{Z}=\begin{bmatrix} K \\ L_2(v_1,v_2,K) \\ L_1(v_1,v_2,K) \\ \beta L_1(v_1,v_2,K)+v_3 \end{bmatrix}.$$

Using $\mathbf{Y}$, receiver 1 can decode all three symbols (u_1,u_2,u_3) and using $\mathbf{Z}$, receiver 2 can decode (v_1,v_2,v_3). Next we prove that the information leakage is only $o(\log P)$.

Security guarantees:

We consider the four slots as a single block and the equivalent channel from $\underline{\mathbf{u}} = (u_1,u_2,u_3)$ to $(\mathbf{Y},\mathbf{H})$ and $(\mathbf{Z},\mathbf{H})$ as a memoryless channel by ignoring the CSI of the previous block. We do the same for the channel from $\underline{\mathbf{v}} = (v_1,v_2,v_3)$ to $(\mathbf{Y},\mathbf{H})$ and $(\mathbf{Z},\mathbf{H})$. Recall that all the random variables $\{u_i,v_i,i=1,2,3\}$ and q are independent and distributed as $\mathcal{CN}(0,P)$.

First, let us consider the confidentiality of the first user's symbols $\underline{\mathbf{u}}$. The information leakage at user 2 is:

$$\begin{aligned}
I(\underline{\mathbf{u}};\mathbf{Z}|\mathbf{H}) =& I(u_1,u_2,u_3;\mathbf{Z}|\mathbf{H}) && (59)\\
=& I(u_1;\mathbf{Z}|\mathbf{H}) && (60)\\
\leq& I(u_1;K|\mathbf{H}) && (61)\\
=& I(u_1;h_{21}(1)u_1+q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp}|\mathbf{H}) && (62)\\
=& h(h_{21}(1)u_1+q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp}|\mathbf{H}) - h(h_{21}(1)u_1+q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp}|u_1,\mathbf{H}) && (63)\\
=& h(h_{21}(1)u_1+q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp}|\mathbf{H}) - h(q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp}|\mathbf{H}) && (64)
\end{aligned}$$

$$= (\log P + o(\log P)) - (\log P + o(\log P)) \tag{65}$$
$$= o(\log P), \tag{66}$$

where (60) follows from the fact that $\mathbf{Z}$ does not have any term involving (u_2, u_3), and (61) follows from the Markov chain $u_1 \to K \to \mathbf{Z}$.

For the second user's symbols, the information leakage at the first receiver is:

$$\begin{aligned} I(\underline{\mathbf{v}}; \mathbf{Y}|\mathbf{H}) =& I(v_1, v_2, v_3; \mathbf{Y}|\mathbf{H}) && (67)\\ =& I(v_1, v_2; \mathbf{Y}|\mathbf{H}) && (68)\\ \leq& I(v_1, v_2; L_1(v_1, v_2, K)|\mathbf{H}) && (69)\\ =& h(L_1(v_1, v_2, K)|\mathbf{H}) - h(L_1(v_1, v_2, K)|v_1, v_2, \mathbf{H}) && (70)\\ \leq& \log P - h(K|v_1, v_2, \mathbf{H}) + o(\log P) && (71)\\ =& \log P - h(K|\mathbf{H}) + o(\log P) && (72)\\ =& \log P - \log P + o(\log P) && (73)\\ =& o(\log P), && (74) \end{aligned}$$

where (68) follows since v_3 does not appear in $\mathbf{Y}$ and (69) follows from the Markov chain $(v_1, v_2) \to L_1(v_1, v_2, K) \to \mathbf{Y}$.

4.3.2 Scheme $S_2^{3/2}$

In this sub-section, we present the scheme $S_2^{3/2}$ which uses the states $(\mathsf{PD}, \mathsf{DP}, \mathsf{PN}, \mathsf{NP})$ with fractions $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ to achieve $(d_1, d_2) = (\frac{3}{4}, \frac{3}{4})$.

Let us consider the utilization of CSIT in the scheme $S_1^{3/2}$ stated above. In the first slot, delayed CSIT is required from the second user, since that knowledge allows the transmitter to reconstruct K and use it in the second slot. Similarly, in the second time slot, delayed CSIT from the first user is required so that the transmitter can reconstruct $L_1(v_1, v_2, K)$ to transmit in the third and fourth slots. However, in the third and fourth slots, the transmitter does not require any CSIT of the first and second users, respectively. Thus, the same scheme works with PN and NP states in the last two slots. Since it is essentially the same scheme interpreted in a different way, the security of the scheme follows from that of $S_1^{3/2}$.

4.4 Schemes Achieving Sum s.d.o.f. of 4/3

4.4.1 Scheme $S_1^{4/3}$

In this sub-section, we present the scheme $S_1^{4/3}$ which uses the states $(\mathsf{PD}, \mathsf{DP}, \mathsf{NN})$ for fractions $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to achieve s.d.o.f. pair $(d_1, d_2) = (\frac{2}{3}, \frac{2}{3})$.

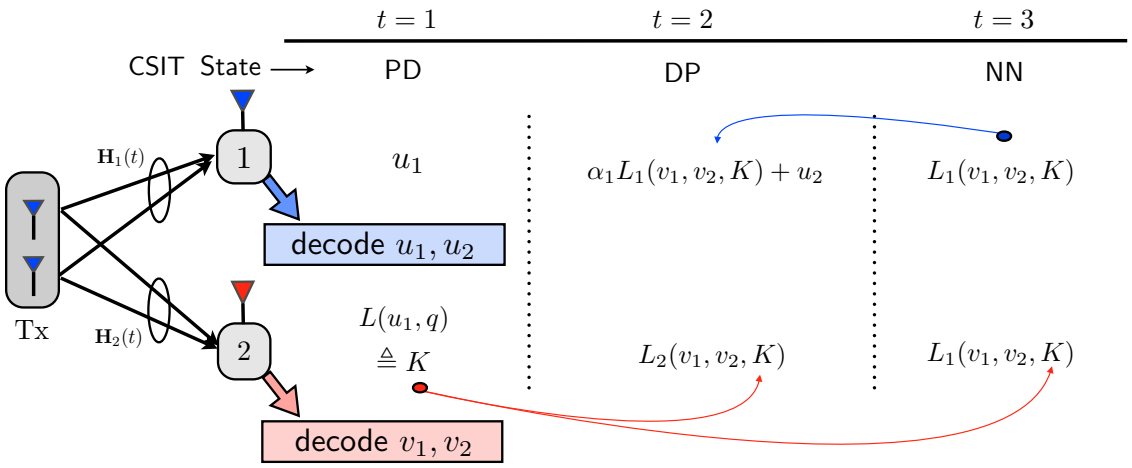


Figure 6: Achieving sum s.d.o.f. of $\frac{4}{3}$ using $S_1^{4/3}$.

We wish to send 2 symbols to each user in 3 time slots. Let (u_1, u_2) and (v_1, v_2) be the symbols intended for the first and second users, respectively. Fig. 6 shows the scheme. It is as follows:

1) At time $t = 1$, $S(1) = \mathsf{PD}$: As the transmitter knows $\mathbf{H}_1(1)$, it sends:

$$\mathbf{X}(1) = [u_1 \quad 0]^T + q\mathbf{H}_1(1)^{\perp}, \tag{75}$$

where $\mathbf{H}_1(1)\mathbf{H}_1(1)^{\perp} = 0$, and q denotes an artificial noise distributed as $\mathcal{CN}(0, P)$. Here $\mathbf{H}_1(1)^{\perp}$ is a 2×1 beamforming vector that ensures that the artificial noise q does not create interference at receiver 1. The receivers' outputs are:

$$Y(1) = h_{11}(1)u_1 \tag{76}$$

$$Z(1) = h_{21}(1)u_1 + q\mathbf{H}_2(1)\mathbf{H}_1(1)^{\perp} \triangleq K. \tag{77}$$

Thus, receiver 1 has observed u_1 while receiver 2 gets a linear combination of u_1 and q, which we denote as K. Due to delayed CSIT from receiver 2, the transmitter can reconstruct K in the next channel use and use it for transmission.

2) At time $t = 2$, $S(2) = \mathsf{DP}$: The transmitter knows $\mathbf{H}_2(2)$ and K. It sends

$$\mathbf{X}(2) = [v_1 + K \quad v_2 + K]^T + u_2\mathbf{H}_2(2)^{\perp}. \tag{78}$$

The received signals are:

$$\begin{aligned} Y(2) &= h_{11}(2)v_1 + h_{12}(2)v_2 + (h_{11}(2) + h_{12}(2))K + u_2\mathbf{H}_1(2)\mathbf{H}_2(2)^{\perp} && (79)\\ &= L_1(v_1, v_2, K) + u_2\mathbf{H}_1(2)\mathbf{H}_2(2)^{\perp} && (80)\\ Z(2) &= h_{21}(2)v_1 + h_{22}(2)v_2 + (h_{21}(2) + h_{22}(2))K \end{aligned}$$

$$\triangleq L_2(v_1, v_2, K), \tag{81}$$

where we have defined $L_1(v_1, v_2, K)$ and $L_2(v_1, v_2, K)$ as independent linear combinations of v_1, v_2 and K at receivers 1 and 2, respectively.

3) At time $t = 3$, $S(3) = \mathsf{NN}$: The transmitter transmits:

$$\mathbf{X}(3) = \begin{bmatrix} L_1(v_1, v_2, K) & 0 \end{bmatrix}^T. \tag{82}$$

The receivers get:

$$Y(3) = h_{11}(3) L_1(v_1, v_2, K) \tag{83}$$

$$Z(3) = h_{21}(3) L_1(v_1, v_2, K). \tag{84}$$

At the end of three slots, therefore, the received outputs can be summarized as:

$$\mathbf{Y} = \begin{bmatrix} u_1 \\ \alpha_1 L_1(v_1, v_2, K) + u_2 \\ L_1(v_1, v_2, K) \end{bmatrix}, \qquad \mathbf{Z} = \begin{bmatrix} K \\ L_2(v_1, v_2, K) \\ L_1(v_1, v_2, K) \end{bmatrix}.$$

Using $\mathbf{Y}$, receiver 1 can decode (u_1, u_2), while receiver 2 can decode (v_1, v_2) using $\mathbf{Z}$. The information leakage is only $o(\log P)$ as we show next.

Security guarantees:

The equivocation calculation follows similar to that of the scheme $S_1^{3/2}$. For the first user's symbols $\underline{\mathbf{u}} = (u_1, u_2)$, we have,

$$I(\underline{\mathbf{u}}; \mathbf{Z}|\mathbf{H}) = I(u_1, u_2; \mathbf{Z}|\mathbf{H}) \tag{85}$$

$$= I(u_1; \mathbf{Z}|\mathbf{H}) \tag{86}$$

$$\leq I(u_1; K|\mathbf{H}) \tag{87}$$

$$= o(\log P), \tag{88}$$

where (86) follows from the fact that $\mathbf{Z}$ does not have any term involving u_2, and (87) follows from the Markov chain $u_1 \to K \to \mathbf{Z}$.

For the second user's symbols, the information leakage at the first receiver is:

$$I(\underline{\mathbf{v}}; \mathbf{Y}|\mathbf{H}) \leq I(v_1, v_2; L_1(v_1, v_2, K)|\mathbf{H}) \tag{89}$$

$$= h(L_1(v_1, v_2, K)|\mathbf{H}) - h(L_1(v_1, v_2, K)|v_1, v_2, \mathbf{H}) \tag{90}$$

$$\leq \log P - h(K|v_1, v_2, \mathbf{H}) + o(\log P) \tag{91}$$

$$= \log P - h(K|\mathbf{H}) + o(\log P) \tag{92}$$

$$= \log P - \log P + o(\log P) \tag{93}$$

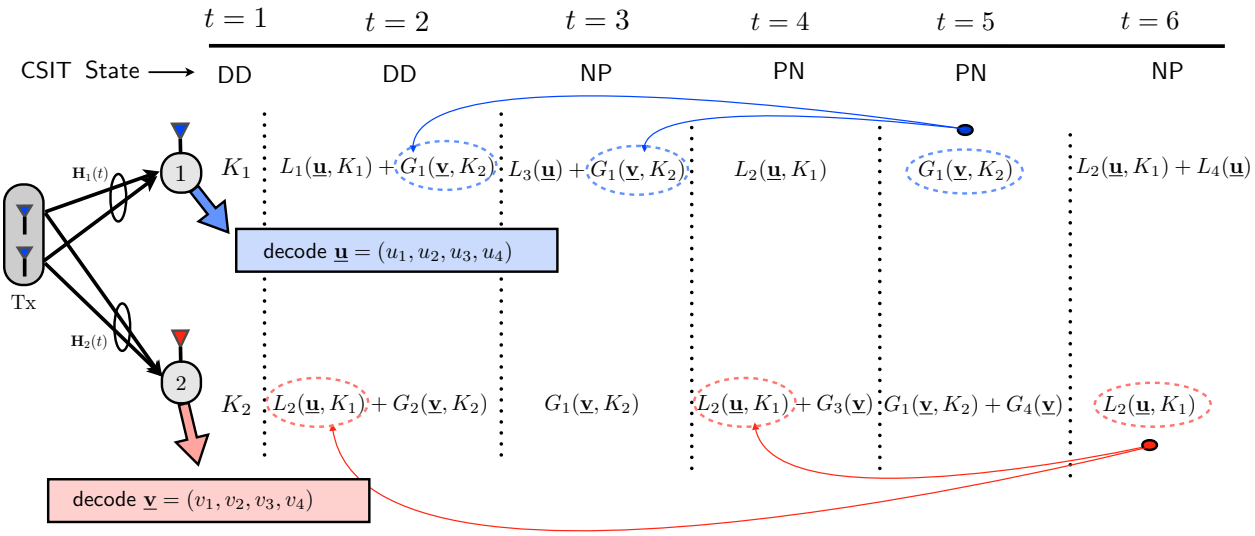


Figure 7: Achieving sum s.d.o.f. $\frac{4}{3}$ using $S_2^{4/3}$.

$$= o(\log P), \tag{94}$$

where (89) follows from the Markov chain $(v_1, v_2) \to L_1(v_1, v_2, K) \to \mathbf{Y}$.

4.4.2 Scheme $S_2^{4/3}$

We now present the scheme $S_2^{4/3}$ which uses the states PN, NP, DD with fractions $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to achieve $(d_1, d_2) = (\frac{2}{3}, \frac{2}{3})$.

In this case we will send 4 symbols to each user in 6 time slots. Let $\underline{\mathbf{u}} = (u_1, u_2, u_3, u_4)$ and $\underline{\mathbf{v}} = (v_1, v_2, v_3, v_4)$ be the symbols intended for the first and second users, respectively. Fig. 7 shows the scheme. It is as follows:

1) At time $t = 1$, $S(1) =$ DD: In this slot, the transmitter sends artificial noise symbols to create keys that can be used in later slots. The channel input is

$$\mathbf{X}(1) = [q_1 \quad q_2]^T, \tag{95}$$

where q_1 and q_2 are i.i.d. as $\mathcal{CN}(0, P)$. The received signals are:

$$Y(1) = h_{11}(1)q_1 + h_{12}(1)q_2 \triangleq K_1 \tag{96}$$

$$Z(1) = h_{21}(1)q_1 + h_{22}(1)q_2 \triangleq K_2. \tag{97}$$

Due to delayed CSIT, the transmitter learns K_1 and K_2 and uses them in the next time slots.

2) At time $t = 2$, $S(2) = \mathsf{DD}$: In this slot, the transmitter sends:

$$\mathbf{X}(2) = [u_1 + u_2 + v_3 + v_4 + K_1 \quad v_1 + v_2 + u_3 + u_4 + K_2]^T . \tag{98}$$

The received signals are:

$$\begin{aligned} Y(2) &= h_{11}(2)(u_1 + u_2 + v_3 + v_4 + K_1) + h_{12}(2)(v_1 + v_2 + u_3 + u_4 + K_2) && (99)\\ &\triangleq L_1(\underline{\mathbf{u}}, K_1) + G_1(\underline{\mathbf{v}}, K_2) && (100)\\ Z(2) &= h_{21}(2)(u_1 + u_2 + v_3 + v_4 + K_1) + h_{22}(2)(v_1 + v_2 + u_3 + u_4 + K_2) && (101)\\ &\triangleq L_2(\underline{\mathbf{u}}, K_1) + G_2(\underline{\mathbf{v}}, K_2). && (102) \end{aligned}$$

Note that since K_1 (or K_2) is known at the first (or second) receiver, it can be removed. The unintended symbols remain buried in the artificial noise, ensuring security. Also, if G_1 (or L_2) could be sent to the second (or first) receiver, it would provide a linear combination of the intended symbols that is linearly independent of G_2 (or L_1). This is what we will do in the third and fourth time slots.

3) At time $t = 3$, $S(3) = \mathsf{NP}$: In this state, the transmitter knows $\mathbf{H}_2$ perfectly. It sends,

$$\mathbf{X}(3) = [G_1(\underline{\mathbf{v}}, K_2) \quad 0]^T + L_3(\underline{\mathbf{u}})\mathbf{H}_2(3)^{\perp}, \tag{103}$$

where L_3 is linearly independent of both L_1 and L_2. The received signals are:

$$\begin{aligned} Y(3) =& h_{11}(3)G_1(\underline{\mathbf{v}}, K_2) + L_3(\underline{\mathbf{u}})\mathbf{H}_1(3)\mathbf{H}_2(3)^{\perp} && (104)\\ Z(3) =& h_{21}(3)G_1(\underline{\mathbf{v}}, K_2). && (105) \end{aligned}$$

4) At time $t = 4$, $S(4) = \mathsf{PN}$: In this state, the transmitter knows $\mathbf{H}_1(4)$ perfectly. It sends,

$$\mathbf{X}(4) = [L_2(\underline{\mathbf{u}}, K_1) \quad 0]^T + G_3(\underline{\mathbf{v}})\mathbf{H}_1(4)^{\perp}, \tag{106}$$

where G_3 is linearly independent of both G_1 and G_2. The received signals are:

$$\begin{aligned} Y(4) =& h_{11}(4)L_2(\underline{\mathbf{u}}, K_1) && (107)\\ Z(4) =& h_{21}(4)L_2(\underline{\mathbf{u}}, K_1) + G_3(\underline{\mathbf{v}})\mathbf{H}_2(4)\mathbf{H}_1(4)^{\perp}. && (108) \end{aligned}$$

Now note that if we could supply G_1 and L_2 to the first and second receivers, respectively, both receivers will end up with 3 linearly independent combinations of their intended symbols. Thus, in the next two slots, the transmitter will supply G_1 and L_2 to the first and second receivers, respectively, as well as send one more linearly independent combination of

the intended information symbols to each receiver.

5) At time $t = 5$, $S(5) = \mathsf{PN}$: In this state, the transmitter knows $\mathbf{H}_1(5)$ perfectly. It sends,

$$\mathbf{X}(5) = [G_1(\underline{\mathbf{v}}, K_2) \quad 0]^T + G_4(\underline{\mathbf{v}})\mathbf{H}_1(5)^{\perp}. \tag{109}$$

The receivers receive:

$$Y(5) = h_{11}(5)G_1(\underline{\mathbf{v}}, K_2) \tag{110}$$
$$Z(5) = h_{21}(5)G_1(\underline{\mathbf{v}}, K_2) + G_4(\underline{\mathbf{v}})\mathbf{H}_2(5)\mathbf{H}_1(5)^{\perp}. \tag{111}$$

6) At time $t = 6$, $S(6) = \mathsf{NP}$: Now the transmitter knows $\mathbf{H}_2(6)$ perfectly, and it sends:

$$\mathbf{X}(6) = [L_2(\underline{\mathbf{u}}, K_1) \quad 0] + L_4(\underline{\mathbf{u}})\mathbf{H}_2(6)^{\perp}. \tag{112}$$

The received signals are:

$$Y(6) = h_{11}(6)L_2(\underline{\mathbf{u}}, K_1) + L_4(\underline{\mathbf{u}})\mathbf{H}_1(6)\mathbf{H}_2(6)^{\perp} \tag{113}$$
$$Z(6) = h_{21}(6)L_2(\underline{\mathbf{u}}, K_1). \tag{114}$$

Let us summarize the received signals at each receiver after these 6 time slots:

$$\mathbf{Y} = \begin{bmatrix} K_1 \\ L_1(\underline{\mathbf{u}}, K_1) + G_1(\underline{\mathbf{v}}, K_2) \\ \alpha_1 G_1(\underline{\mathbf{v}}, K_2) + L_3(\underline{\mathbf{u}}) \\ L_2(\underline{\mathbf{u}}, K_1) \\ G_1(\underline{\mathbf{v}}, K_2) \\ \alpha_2 L_2(\underline{\mathbf{u}}, K_1) + L_4(\underline{\mathbf{u}}) \end{bmatrix}, \qquad \mathbf{Z} = \begin{bmatrix} K_2 \\ L_2(\underline{\mathbf{u}}, K_1) + G_2(\underline{\mathbf{v}}, K_2) \\ G_1(\underline{\mathbf{v}}, K_2) \\ \beta_1 L_2(\underline{\mathbf{u}}, K_1) + G_3(\underline{\mathbf{v}}) \\ \beta_2 G_1(\underline{\mathbf{v}}, K_2) + G_4(\underline{\mathbf{v}}) \\ L_2(\underline{\mathbf{u}}, K_1) \end{bmatrix}.$$

The information symbols can now be decoded at the intended receivers from these observations. Also the leakage of information is only $o(\log P)$, as we prove next.

Security guarantees:

For the first user's symbols $\underline{\mathbf{u}} = (u_1, u_2, u_3, u_4)$, we have,

$$I(\underline{\mathbf{u}}; \mathbf{Z}|\mathbf{H}) \leq I(\underline{\mathbf{u}}; L_2(\underline{\mathbf{u}}, K_1)|\mathbf{H}) \tag{115}$$
$$= h(L_2(\underline{\mathbf{u}}, K_1)|\mathbf{H}) - h(L_2(\underline{\mathbf{u}}, K_1)|\underline{\mathbf{u}}, \mathbf{H}) \tag{116}$$
$$\leq \log P - h(K_1|\underline{\mathbf{u}}, \mathbf{H}) + o(\log P) \tag{117}$$
$$= \log P - h(K_1|\mathbf{H}) + o(\log P) \tag{118}$$
$$= \log P - \log P + o(\log P) \tag{119}$$

$$= o(\log P), \tag{120}$$

where (115) follows from the Markov chain $U \rightarrow L_2(\underline{\mathbf{u}}, K_1) \rightarrow \mathbf{Z}$.

For the second user's symbols, the information leakage at the first receiver is:

$$\begin{aligned} I(\underline{\mathbf{v}}; \mathbf{Y}|\mathbf{H}) \leq & I(\underline{\mathbf{v}}; G_1(\underline{\mathbf{v}}, K_2)|\mathbf{H}) && (121)\\ = & h(G_1(\underline{\mathbf{v}}, K_2)|\mathbf{H}) - h(G_1(\underline{\mathbf{v}}, K_2)|\underline{\mathbf{v}}, \mathbf{H}) && (122)\\ \leq & \log P - h(K_2|\underline{\mathbf{v}}, \mathbf{H}) + o(\log P) && (123)\\ = & \log P - h(K_2|\mathbf{H}) + o(\log P) && (124)\\ = & \log P - \log P + o(\log P) && (125)\\ = & o(\log P), && (126) \end{aligned}$$

where (89) follows from the Markov chain $\underline{\mathbf{v}} \rightarrow G_1(\underline{\mathbf{v}}, K_2) \rightarrow \mathbf{Y}$.

4.5 Schemes Achieving Sum s.d.o.f. of 1

4.5.1 Scheme S_1^1

We first recap the scheme S_1^1 which uses the state DD to achieve $(d_1, d_2) = (\frac{1}{2}, \frac{1}{2})$. This scheme was presented in [43]. The scheme was used to transmit 2 information symbols to each receiver in 4 time slots. At $t = 1$, the transmitter sends artificial noise symbols using both antennas. The received signals act as keys K_1 and K_2 for the respective users 1 and 2. Since there is delayed CSIT, the transmitter can reconstruct these keys and use them in the next slots. At $t = 2$, the transmitter sends the two information symbols (u_1, u_2) intended for the first receiver linearly combined with the first user's key. Thus, the first user can retrieve a linear combination of just its intended symbols. However, the second user gets a linear combination $L(u_1, u_2, K_1)$. Due to delayed CSIT however, the transmitter can reconstruct L. In the third slot, the roles of the receivers are reversed and the transmitter sends the second user's symbols (v_1, v_2) linearly combined with the second user's key K_2. This allows the second user to retrieve a linear combination of just its information symbol, which however remain secure at the first user, which receives $G(v_1, v_2, K_2)$. In the fourth slot, the transmitter sends a linear combination of L and G. Essentially this provides the first user with L, from which it can eliminate K_1 to get another independent linear combination of (u_1, u_2). A similar situation takes place at the second user. Finally, each user has two linearly independent combinations of two symbols and thus can decode the information symbols intended for it. The information leakage is only $o(\log P)$, as shown in [43].

4.5.2 Scheme S_2^1

In this sub-section, we present the scheme S_2^1 which uses the states $(\mathsf{DD}, \mathsf{NN})$ with fractions $(\frac{1}{2}, \frac{1}{2})$ to achieve $(d_1, d_2) = (\frac{1}{2}, \frac{1}{2})$.

The scheme S_1^1 requires delayed CSIT from at least one user for the first 3 time slots. We need to modify this scheme to ensure that delayed CSIT is required only for 2 of the 4 time slots. Fig. 8 shows the new scheme. It is as follows:

1) At time $t = 1$, $S(1) = \mathsf{DD}$: The strategy in this slot is the same as in the scheme S_1^1. In this slot, the transmitter sends artificial noise symbols to create keys that can be used in later slots. The channel input is

$$\mathbf{X}(1) = \begin{bmatrix} q_1 & q_2 \end{bmatrix}^T, \tag{127}$$

where q_1 and q_2 are i.i.d. as $\mathcal{CN}(0, P)$. The received signals are:

$$Y(1) = h_{11}(1)q_1 + h_{12}(1)q_2 \triangleq K_1 \tag{128}$$

$$Z(1) = h_{21}(1)q_1 + h_{22}(1)q_2 \triangleq K_2. \tag{129}$$

Due to delayed CSIT, the transmitter learns K_1 and K_2 and uses them in the next time slots.

2) At time $t = 2$, $S(2) = \mathsf{DD}$: Instead of sending only the first user's symbols as in scheme S_1^1, the transmitter now sends linear combination of both users' symbols. It sends:

$$\mathbf{X}(2) = \begin{bmatrix} u_1 + v_1 + K_1 & u_2 + v_2 + K_2 \end{bmatrix}^T. \tag{130}$$

The received signals are:

$$Y(2) = h_{11}(u_1 + v_1 + K_1) + h_{12}(u_2 + v_2 + K_2) \tag{131}$$

$$\triangleq L_1(u_1, u_2, K_1) + G_1(v_1, v_2, K_2) \tag{132}$$

$$Z(2) = h_{21}(u_1 + v_1 + K_1) + h_{22}(u_2 + v_2 + K_2) \tag{133}$$

$$\triangleq L_2(u_1, u_2, K_1) + G_2(v_1, v_2, K_2). \tag{134}$$

We notice that if L_2 and G_1 could be provided to both users, each user can get 2 linear combinations of the symbols intended for it and hence decode both symbols. Hence, in the remaining two slots, we will transmit L_2 and G_1 to both users and this will not require any CSIT from any user.

3) At time $t = 3$, $S(3) = \mathsf{NN}$: The transmitter does not have any CSIT. It sends:

$$\mathbf{X}(3) = \begin{bmatrix} L_2(u_1, u_2, K_1) & 0 \end{bmatrix}^T. \tag{135}$$

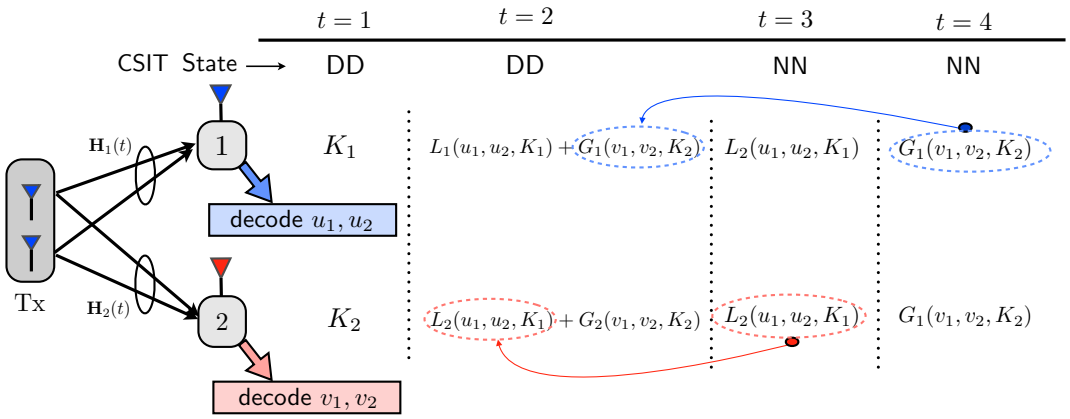


Figure 8: Achieving sum s.d.o.f. of 1 using S_2^1.

The received signals are:

$$Y(3) = h_{11}(3)L_2(u_1,u_2,K_1) \tag{136}$$
$$Z(3) = h_{21}(3)L_2(u_1,u_2,K_1). \tag{137}$$

4) At time $t=4$, $S(4)=\mathsf{NN}$: The transmitter sends:

$$\mathbf{X}(4) = \begin{bmatrix} G_1(v_1,v_2,K_2) & 0 \end{bmatrix}^T. \tag{138}$$

The received signals are:

$$Y(4) = h_{11}(4)G_1(v_1,v_2,K_2) \tag{139}$$
$$Z(4) = h_{21}(4)G_1(v_1,v_2,K_2). \tag{140}$$

Thus, at the end of 4 slots the received signals may be summarized as:

$$\mathbf{Y} = \begin{bmatrix} K_1 \\ L_1(u_1,u_2,K_1)+G_1(v_1,v_2,K_2) \\ L_2(u_1,u_2,K_1) \\ G_1(v_1,v_2,K_2) \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} K_2 \\ L_2(u_1,u_2,K_1)+G_2(v_1,v_2,K_2) \\ L_2(u_1,u_2,K_1) \\ G_1(v_1,v_2,K_2) \end{bmatrix}.$$

Clearly, user 1 can decode (u_1,u_2) and user 2 can get (v_1,v_2). The information leakage is at most $o(\log P)$ as we show below.

Security guarantees:

For the first user's symbols $\underline{\mathbf{u}} = (u_1,u_2)$, we have,

$$I(\underline{\mathbf{u}};\mathbf{Z}|\mathbf{H}) \leq I(\underline{\mathbf{u}};L_2(\underline{\mathbf{u}},K_1)|\mathbf{H}) \tag{141}$$
$$= h(L_2(\underline{\mathbf{u}},K_1)|\mathbf{H}) - h(L_2(\underline{\mathbf{u}},K_1)|\underline{\mathbf{u}},\mathbf{H}) \tag{142}$$

$$\leq \log P - h(K_1|\underline{\mathbf{u}}, \mathbf{H}) + o(\log P) \tag{143}$$
$$= \log P - h(K_1|\mathbf{H}) + o(\log P) \tag{144}$$
$$= \log P - \log P + o(\log P) \tag{145}$$
$$= o(\log P), \tag{146}$$

where (141) follows from the Markov chain $U \to L_2(\underline{\mathbf{u}}, K_1) \to \mathbf{Z}$.

For the second user's symbols $\underline{\mathbf{v}} = (v_1, v_2)$, the information leakage at the first receiver is:

$$I(\underline{\mathbf{v}}; \mathbf{Y}|\mathbf{H}) \leq I(\underline{\mathbf{v}}; G_1(\underline{\mathbf{v}}, K_2)|\mathbf{H}) \tag{147}$$
$$= h(G_1(\underline{\mathbf{v}}, K_2)|\mathbf{H}) - h(G_1(\underline{\mathbf{v}}, K_2)|\underline{\mathbf{v}}, \mathbf{H}) \tag{148}$$
$$\leq \log P - h(K_2|\underline{\mathbf{v}}, \mathbf{H}) + o(\log P) \tag{149}$$
$$= \log P - h(K_2|\mathbf{H}) + o(\log P) \tag{150}$$
$$= \log P - \log P + o(\log P) \tag{151}$$
$$= o(\log P), \tag{152}$$

where (147) follows from the Markov chain $\underline{\mathbf{v}} \to G_1(\underline{\mathbf{v}}, K_2) \to \mathbf{Y}$.

4.5.3 Scheme S_3^1

We next present a novel scheme S_3^1 which uses the states $(\mathsf{DN}, \mathsf{ND})$ with fractions $(\frac{1}{2}, \frac{1}{2})$ to achieve $(d_1, d_2) = (\frac{1}{2}, \frac{1}{2})$. In particular, we present a scheme which achieves the s.d.o.f. pair $(d_1, d_2) = \left(\frac{2n}{4n+1}, \frac{2n}{4n+1}\right)$ as a function of the block length n. Taking the limit $n \to \infty$ yields the s.d.o.f. pair $\left(\frac{1}{2}, \frac{1}{2}\right)$.

The scheme is shown in Fig. 9. Unlike all the other schemes in this paper where the optimal sum s.d.o.f. can be achieved within a finite number of time slots, this scheme cannot achieve sum s.d.o.f. of 1 in a finite number of slots. Indeed, there does not exist a scheme that can achieve sum s.d.o.f. of 1 in finitely many slots. To see why, assume that there exists such a scheme with n slots. In this scheme, states DN and ND occur for equal fractions of time; thus, $\lambda_D = \lambda_N = \frac{1}{2}$. Now, note that the delayed CSIT in the last slot cannot be used; thus, the scheme would work equally well if the last slot were NN instead of DN or ND. However, changing the state in the last slot to NN would imply $\lambda_D < \frac{1}{2}$, which in turn implies that $d_1 + d_2 < 1$ from (18). Thus, no scheme that uses only a finite number of slots can achieve a sum s.d.o.f. of 1.

Here we provide an asymptotic scheme that achieves a sum s.d.o.f. of $\frac{4n}{4n+1}$ in n slots. As the number of slots $n \to \infty$, the sum s.d.o.f. approaches 1. We wish to send $2n$ symbols to each receiver in $4n+1$ time slots. The scheme involves transmission in 4 blocks where the first 3 blocks, say A, B and C each have n time slots, while the last block D has $n+1$ slots;

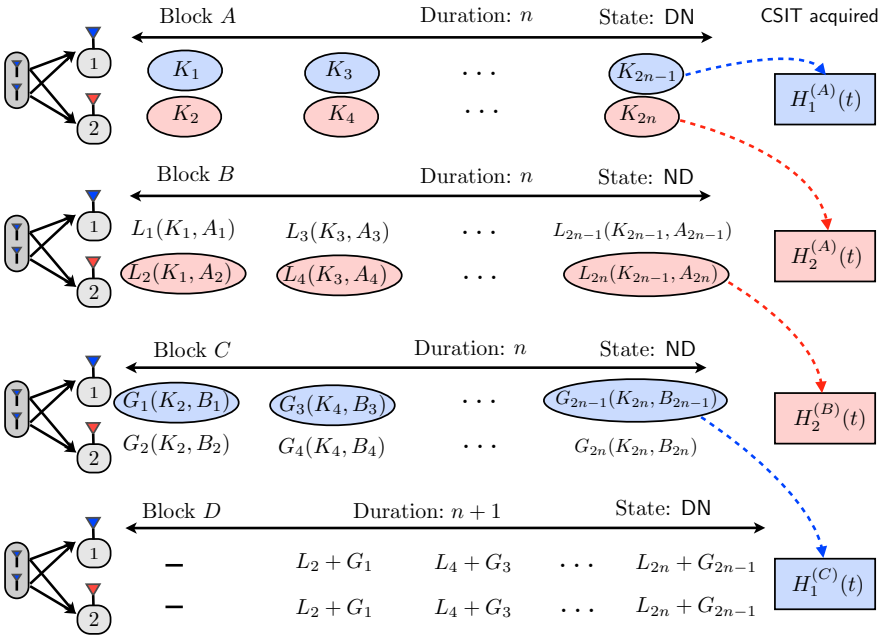


Figure 9: Achieving sum s.d.o.f. of $4n/(4n+1)$ using scheme S_3^1.

thus, a total of $4n+1$ time slots are required in the scheme. The scheme is as follows:

1) In block A, $S(t) = \mathsf{DN}$: In each time slot i in block A, the transmitter generates two artificial noise symbols and sends them using its two antennas. The receivers receive different linear combinations of the two artificial noise symbols K_{2i-1} and K_{2i} as shown in Fig. 9. Due to delayed CSIT from the first user, the transmitter can reconstruct each of $K_{2i-1}, i = 1, \ldots,$ by the end of block A. Thus, they can act as shared keys between the transmitter and the first receiver. However, since the second receiver does not feedback any CSIT (due to the fact that the state in the block is DN), the transmitter cannot reconstruct the observations of the second receiver at the end of block A.

2) In block B, $S(t) = \mathsf{ND}$: At the beginning of this slot, the transmitter has the keys $K_{2i-1}, i = 1, \ldots, n$ shared with the first user. It uses these keys to send information intended for the first user. It creates $2n$ linearly independent combinations of the $2n$ symbols intended for the first receiver: $a_1, \ldots, a_{2n}$. In slot i, it transmits

$$\mathbf{X}^B(i) = \begin{bmatrix} a_{2i-1} + K_{2i-1} & a_{2i} + K_{2i-1} \end{bmatrix}^T. \tag{153}$$

The first and second receivers receive linearly independent combinations $L_{2i-1}(A_{2i-1}, K_{2i-1})$ and $L_{2i}(A_{2i}, K_{2i-1})$ in slot i, where A_i denotes the ith linear combination of the first user's symbols, as shown in Fig. 9. Since the state is ND, the second user provides delayed CSIT to the transmitter. In the ith slot, the second user feeds back $\mathbf{H}_2^A(i)$, that is, the channel coefficients of the second user in slot i within block A. *Note that this is unlike any other achievable scheme we have encountered so far; in all other schemes, the receiver feeds back*

the channel coefficients of the current slot which appears as delayed CSIT at the beginning of the next slot. Thus, at the end of slot B, the transmitter has all the channel coefficients of the second user from block A; thus, it can reconstruct the outputs of the second receiver in block A, $K_{2i}, i = 1, \ldots, n$, which now act as shared keys between the transmitter and the second receiver.

3) In block C, $S(t) = \mathsf{ND}$: At the beginning of this slot, the transmitter has the keys $K_{2i}, i = 1, \ldots, n$ shared with the second user. It uses these keys to send information securely to the second user. It creates $2n$ linearly independent combinations of the $2n$ symbols intended for the second receiver: $b_1, \ldots, a_{2n}$. In slot i, it transmits

$$\mathbf{X}^C(i) = \begin{bmatrix} b_{2i-1} + K_{2i-1} & b_{2i} + K_{2i-1} \end{bmatrix}^T. \tag{154}$$

The first and second receivers receive linearly independent combinations $G_{2i-1}(B_{2i-1}, K_{2i})$ and $G_{2i}(B_{2i}, K_{2i})$ in slot i, where B_i denotes the ith linear combination of the second user's symbols, as shown in Fig. 9. As CSIT, in the ith slot, the second user feeds back the channel coefficients $\mathbf{H}_2^B(i)$, which allows the transmitter to reconstruct $L_{2i}(A_{2i}, K_{2i-1})$. Note that now if $L_{2i}(A_{2i}, K_{2i-1})$ and $G_{2i-1}(B_{2i-1}, K_{2i})$ could be exchanged, each of the receivers would receive $2n$ linear combinations of the $2n$ symbols intended for it, thus, allowing both receivers to decode their own messages. However, $G_{2i-1}(B_{2i-1}, K_{2i})$ is not known to the transmitter yet, since the first user has not fed back its channel in block C. This CSIT will be obtained in the next block.

4) In block D, $S(t) = \mathsf{ND}$: The transmitter wishes to send the symbols $L_{2i}(A_{2i}, K_{2i-1}) + G_{2i-1}(B_{2i-1}, K_{2i}), i = 1, \ldots, n$, in this block. To do so, the transmitter does not transmit anything in the first slot in this block. It only acquires the channel coefficients $\mathbf{H}_1^C(i)$ from the first user who is supplying delayed CSIT in this block. In the ith slot, $i = 1, \ldots, n$, the transmitter acquires the channel coefficients $\mathbf{H}_1^C(i)$ and transmits:

$$\mathbf{X}^D(i) = \begin{bmatrix} L_{2i-2}(A_{2i-2}, K_{2i-3}) + G_{2i-3}(B_{2i-3}, K_{2i-2}) & 0 \end{bmatrix}^T, \quad i = 2, \ldots, n+1. \tag{155}$$

The first user can now obtain $L_{2i-1}(A_{2i-1}, K_{2i-1})$ and $L_{2i}(A_{2i}, K_{2i-1})$ for every $i = 1, \ldots, n$, while the second user obtains $G_{2i-1}(B_{2i-1}, K_{2i})$ and $G_{2i}(B_{2i}, K_{2i})$ for $i = 1, \ldots, n$. Now by eliminating the respective keys, each user can decode the $2n$ symbols intended for it from the $2n$ linearly independent combinations available to it. Also the keys ensure the confidentiality, and the information leakage is only $o(\log P)$, as we show next.

Security guarantees:

Let $\underline{\mathbf{u}} = (a_1, \ldots, a_{2n})$ and $\underline{\mathbf{v}} = (b_1, \ldots, b_{2n})$ be the symbols intended for users 1 and 2, respectively. The leakage of $\underline{\mathbf{u}}$ at user 2 is given by

$$I(\underline{\mathbf{u}}; \mathbf{Z}|\mathbf{H}) \leq I(\underline{\mathbf{u}}; \{L_{2i}(A_{2i}, K_{2i-1})\}_{i=1}^n |\mathbf{H}) \tag{156}$$

$$
\begin{aligned}
&= h(\{L_{2i}(A_{2i}, K_{2i-1})\}_{i=1}^{n} | \mathbf{H}) - h(\{L_{2i}(A_{2i}, K_{2i-1})\}_{i=1}^{n} | \underline{\mathbf{u}}, \mathbf{H}) && (157)\\
&\leq n \log P - h(\{K_{2i-1}\}_{i=1}^{n} | \mathbf{H}) + o(\log P) && (158)\\
&= n \log P - n \log P + o(\log P) && (159)\\
&= o(\log P), && (160)
\end{aligned}
$$

where (156) follows due to the Markov chain $\underline{\mathbf{u}} \rightarrow \{L_{2i}(A_{2i}, K_{2i-1})\}_{i=1}^{n} \rightarrow \mathbf{Z}$, and (159) follows from the fact that $\{K_{2i-1}\}_{i=1}^{n}$ are mutually independent and each is distributed as $\mathcal{N}(0, P)$.

Similarly, for the second user's symbols, the leakage at the first user is given by,

$$
\begin{aligned}
I(\underline{\mathbf{v}}; \mathbf{Y} | \mathbf{H}) &\leq I(\underline{\mathbf{v}}; \{G_{2i-1}(B_{2i-1}, K_{2i})\}_{i=1}^{n} | \mathbf{H}) && (161)\\
&= h(\{G_{2i-1}(B_{2i-1}, K_{2i})\}_{i=1}^{n} | \mathbf{H}) - h(\{G_{2i-1}(B_{2i-1}, K_{2i})\}_{i=1}^{n} | \underline{\mathbf{v}}, \mathbf{H}) && (162)\\
&\leq n \log P - h(\{K_{2i}\}_{i=1}^{n} | \mathbf{H}) + o(\log P) && (163)\\
&= n \log P - n \log P + o(\log P) && (164)\\
&= o(\log P), && (165)
\end{aligned}
$$

where (161) follows due to the Markov chain $\underline{\mathbf{v}} \rightarrow \{G_{2i-1}(B_{2i-1}, K_{2i})\}_{i=1}^{n} \rightarrow \mathbf{Y}$, and (164) follows from the fact that $\{K_{2i}\}_{i=1}^{n}$ are mutually independent and each is distributed as $\mathcal{N}(0, P)$.

4.6 Schemes Achieving Sum s.d.o.f. of 2/3

4.6.1 Scheme $S_1^{2/3}$

The scheme $S_1^{2/3}$ uses the state DD to achieve $(d_1, d_2) = (\frac{2}{3}, 0)$. Such a scheme was presented in [43]. The scheme can be summarized as follows. At time $t = 1$, the transmitter sends two artificial noise symbols using its two antennas. Each user receives a different linear combination of the noise symbols and they act as keys. Let K_1 and K_2 be the keys at receivers 1 and 2, respectively. Due to delayed CSIT, the transmitter can reconstruct K_1. At time $t = 2$, the transmitter sends the two symbols intended for the first receiver (u_1, u_2), linearly combined with K_1. Receiver 1 can remove K_1 from its received signal and get one linear combination of (u_1, u_2) at the end of this slot. The second user receives a linear combination of u_1, u_2 and K_1, say $L(u_1, u_2, K_1)$; however, not knowing K_1, it cannot decode the information symbols. Due to delayed CSIT, the transmitter learns L and transmits it in $t = 3$. The second receiver gets no new information but the first receiver can get a second linear combination of (u_1, u_2) by eliminating K_1 from L. This allows receiver 1 to decode (u_1, u_2), while the information leakage to receiver 2 is $o(\log P)$.

4.6.2 Scheme $S_2^{2/3}$

The scheme $S_2^{2/3}$ uses the states $(\mathsf{DD}, \mathsf{NN})$ with fractions $(\frac{2}{3}, \frac{1}{3})$ to achieve $(d_1, d_2) = (\frac{2}{3}, 0)$. We note that in scheme $S_1^{2/3}$, the delayed CSIT in slot $t = 3$ is not required. Thus, the scheme can work with the states $(\mathsf{DD}, \mathsf{NN})$ with fractions $(\frac{2}{3}, \frac{1}{3})$, and we call this $S_2^{2/3}$.

4.6.3 Scheme $S_3^{2/3}$

Finally, the scheme $S_3^{2/3}$ uses the states $(\mathsf{DN}, \mathsf{ND}, \mathsf{NN})$ with fractions $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to achieve $(d_1, d_2) = (\frac{2}{3}, 0)$. We notice that instead of having DD state in the first two slots, it suffices to have DN in the first slot (since the transmitter does not need K_2) and ND in the second slot (since the transmitter only needs to reconstruct the second user's received signal L). Thus, it suffices to have the states $(\mathsf{DN}, \mathsf{ND}, \mathsf{NN})$ with fractions $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for the scheme to work, and we call this $S_3^{2/3}$.

5 Achievability

Now that we have all the required constituent schemes summarized in Table 1, we proceed to show how these schemes can be combined to achieve the region stated in Theorem 1. We restate the region of Theorem 1 here for convenience:

$$d_1 \leq \min\left(\frac{2 + 2\lambda_P - \lambda_{PP}}{3}, 1 - \lambda_{NN}\right) \tag{166}$$

$$d_2 \leq \min\left(\frac{2 + 2\lambda_P - \lambda_{PP}}{3}, 1 - \lambda_{NN}\right) \tag{167}$$

$$3d_1 + d_2 \leq 2 + 2\lambda_P \tag{168}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_P \tag{169}$$

$$d_1 + d_2 \leq 2(\lambda_P + \lambda_D). \tag{170}$$

We classify this region into two cases:

- Case A: in which $d_1 + d_2$ bound of (170) is inactive. This corresponds to the condition

$$1 + \lambda_P \leq 2\lambda_P + 2\lambda_D, \tag{171}$$

  which is equivalent to

$$\lambda_N \leq \lambda_D. \tag{172}$$

- Case B: in which $d_1 + d_2$ bound of (170) is active which corresponds to

$$\lambda_N > \lambda_D. \tag{173}$$

In the next two sub-sections, we present the achievability for each of these cases separately.

5.1 Achievability for Case A: $\lambda_D \geq \lambda_N$

For Case A, the s.d.o.f. region reduces to:

$$d_1 \leq \min\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}, 1-\lambda_{NN}\right) \tag{174}$$

$$d_2 \leq \min\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}, 1-\lambda_{NN}\right) \tag{175}$$

$$3d_1 + d_2 \leq 2 + 2\lambda_P \tag{176}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_P. \tag{177}$$

Depending on which single user bound is active, we consider two cases:

1. $\frac{2+2\lambda_P-\lambda_{PP}}{3} \leq 1-\lambda_{NN}$, which is equivalent to the condition $\lambda_{DD} + 2\lambda_{DN} \geq 2\lambda_{NN}$,
2. $\frac{2+2\lambda_P-\lambda_{PP}}{3} \geq 1-\lambda_{NN}$, which is equivalent to the condition $\lambda_{DD} + 2\lambda_{DN} \leq 2\lambda_{NN}$.

As shown in Fig. 10, due to symmetry, it suffices to achieve the points P_1 and P_2 in each case.

5.1.1 Achievability of Point P_1

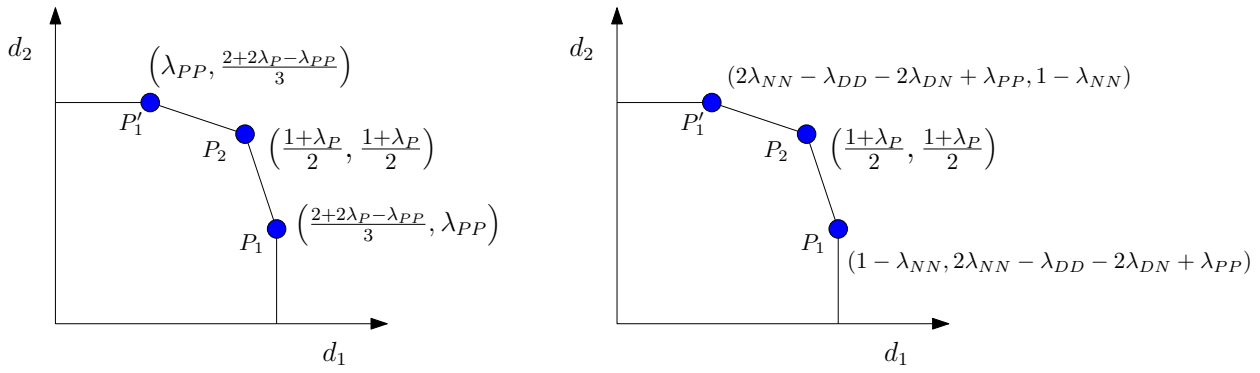
We first show the achievability of the point P_1 in both cases. To do so, let us consider the two cases one by one:

1. $\lambda_{DD} + 2\lambda_{DN} \geq 2\lambda_{NN}$: In this case, the single user bounds are:

$$d_1 \leq \frac{2+2\lambda_P-\lambda_{PP}}{3} \tag{178}$$

$$d_2 \leq \frac{2+2\lambda_P-\lambda_{PP}}{3}. \tag{179}$$

As seen in Fig. 10a, the point P_1 is $\left(\frac{2+2\lambda_{\mathsf{P}}-\lambda_{\mathsf{PP}}}{3}, \lambda_{\mathsf{PP}}\right)$. To achieve this point, using the state PP, we achieve $(1,1)$, with $\mathsf{PD}, \mathsf{DP}, \mathsf{PN}, \mathsf{NP}$, we achieve the pair $(1,0)$ either through zero-forcing, or by transmitting artificial noise in a direction orthogonal to the first user's channel. For the states $(\mathsf{DD}, \mathsf{NN}) \sim (\frac{2}{3}, \frac{1}{3})$, and $(\mathsf{DN}, \mathsf{ND}, \mathsf{NN}) \sim (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$,



(a) $\lambda_{DD}+2\lambda_{DN} \geq 2\lambda_{NN}$. (b) $\lambda_{DD}+2\lambda_{DN} \leq 2\lambda_{NN}$.

Figure 10: s.d.o.f. regions in case A.

we achieve the pair $(\frac{2}{3}, 0)$ by using the schemes $S_2^{2/3}$ and $S_3^{2/3}$, respectively. Essentially, the NN state can be fully alternated with the DD state and the DN and ND states to achieve $\frac{2}{3}$ s.d.o.f. for user 1.

Time sharing yields the following s.d.o.f. pair:

$$d_2 = \lambda_{PP} \tag{180}$$

$$d_1 = \lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_2^{2/3}}(\lambda_{DD} + 2\lambda_{DN} + \lambda_{NN}) \tag{181}$$

$$= 2\lambda_P - \lambda_{PP} + \frac{2}{3}(\lambda_{DD} + \lambda_{NN}) \tag{182}$$

$$= 2\lambda_P - \lambda_{PP} + \frac{2}{3}(1 - 2\lambda_P + \lambda_{PP}) \tag{183}$$

$$= \frac{2 + 2\lambda_P - \lambda_{PP}}{3}. \tag{184}$$

2. $\lambda_{DD}+2\lambda_{DN} \leq 2\lambda_{NN}$: In this case the single user bounds are:

$$d_1 \leq 1 - \lambda_{NN} \tag{185}$$

$$d_2 \leq 1 - \lambda_{NN}. \tag{186}$$

Again, we wish to achieve the point P_1 in Fig. 10b. The point P_1 is given by:

$$P_1 : (d_1, d_2) = (1 - \lambda_{NN}, \lambda_{PP} + (2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD})). \tag{187}$$

Here we consider two further subcases

- $\lambda_{NN} \leq \lambda_{DD} + \lambda_{DN}$: In this case, to achieve the point P_1, we first use up the full DN and ND states with a part of the NN state using scheme $S_3^{2/3}$. We alternate

the remaining $(\lambda_{NN} - \lambda_{DN})$ duration of NN state with the DD state using two schemes: $S_2^{2/3}$ and S_2^1. Note that in this subcase, $0 \leq 2(\lambda_{DD}+\lambda_{DN}-\lambda_{NN}) \leq \lambda_{DD}$. We use the state DD for duration $2(\lambda_{DD}+\lambda_{DN}-\lambda_{NN})$ and state NN for duration $(\lambda_{DD}+\lambda_{DN}-\lambda_{NN})$ together using scheme $S_2^{2/3}$ to achieve the s.d.o.f. pair $\left(\frac{2}{3}, 0\right)$. The remaining $(2\lambda_{NN}-2\lambda_{DN}-\lambda_{DD})$ duration of the state NN is alternated with the remaining $(2\lambda_{NN}-2\lambda_{DN}-\lambda_{DD})$ duration of state DD using the scheme S_2^1 to achieve the s.d.o.f. pair $\left(\frac{1}{2}, \frac{1}{2}\right)$. The state PP allows us to achieve the s.d.o.f. pair $(1, 1)$ while the remaining states PD, DP, PN, and NP each achieves $(1, 0)$. Thus, by using time sharing, the s.d.o.f. pair is:

$$\begin{aligned} d_1 =& \lambda_{PP} + 1 \times 2\lambda_{PD} + 1 \times 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3\lambda_{DN} + \underbrace{\frac{2}{3}}_{S_2^{2/3}} \times 3(\lambda_{DD} + \lambda_{DN} - \lambda_{NN}) \\ &+ \underbrace{\frac{1}{2}}_{S_2^1} \times 2(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}) \qquad (188) \\ =& 1 - \lambda_{NN} \qquad (189) \\ d_2 =& \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}) \\ =& \lambda_{PP} + (2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}), \qquad (190) \end{aligned}$$

which is precisely the point P_1.

- $\lambda_{NN} \geq \lambda_{DD} + \lambda_{DN}$: In this case, the state NN cannot be completely used with the states DD, DN and ND. But we note that $\lambda_D \geq \lambda_N$ implies that $\lambda_D \geq \lambda_{NN}$. We first use up the DN and ND states by alternating with the NN state using scheme $S_3^{2/3}$. A portion λ_{DD} of the remaining $(\lambda_{NN} - \lambda_{DN})$ duration of the NN state uses up the DD state in scheme S_2^1 achieving the pair $\left(\frac{1}{2}, \frac{1}{2}\right)$. The remaining $(\lambda_{NN} - \lambda_{DN} - \lambda_{DD})$ portion of the NN state is used with the PD and DP states through the scheme $S_1^{4/3}$ to achieve the pair $\left(\frac{2}{3}, \frac{2}{3}\right)$. For the remainder of the state PD, DP and the states PN, NP, we can achieve the pair $(1, 0)$, while $(1, 1)$ is achieved in the PP state. By time sharing, we get

$$\begin{aligned} d_1 =& \lambda_{PP} + 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3\lambda_{DN} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3(\lambda_{NN} - \lambda_{DN} - \lambda_{DD}) \\ &+ 2(\lambda_{PD} - \lambda_{NN} + \lambda_{DN} + \lambda_{DD}) + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} \qquad (191) \end{aligned}$$

$$=1-\lambda_{NN} \tag{192}$$

$$d_2 =\lambda_{PP}+\underbrace{\frac{2}{3}}_{S_1^{4/3}}\times 3(\lambda_{NN}-\lambda_{DN}-\lambda_{DD})+\underbrace{\frac{1}{2}}_{S_2^1}\times 2\lambda_{DD} \tag{193}$$

$$=\lambda_{PP}+2\lambda_{NN}-2\lambda_{DN}-\lambda_{DD}, \tag{194}$$

which is again the point P_1.

5.1.2 Achieving the Sum s.d.o.f. Achieving Point P_2

The point P_2 corresponds to:

$$P_2 : (d_1, d_2) = \left(\frac{1+\lambda_P}{2}, \frac{1+\lambda_P}{2}\right). \tag{195}$$

We rewrite the condition $\lambda_D \geq \lambda_N$ corresponding to case A as:

$$\lambda_{PD}+\lambda_{DD} \geq \lambda_{PN}+\lambda_{NN}. \tag{196}$$

From this condition it is not immediately clear how the constituent schemes should be jointly utilized. Hence we break this condition into three mutually exclusive cases:

1. Sub-case $A1$: $\lambda_{PD} \geq \lambda_{PN}$ and $\lambda_{DD} \geq \lambda_{NN}$,
2. Sub-case $A2$: $\lambda_{PD} \geq \lambda_{PN}$ and $\lambda_{DD} \leq \lambda_{NN}$,
3. Sub-case $A3$: $\lambda_{PD} \leq \lambda_{PN}$ and $\lambda_{DD} \geq \lambda_{NN}$.

Now, we consider these three sub-cases one by one:

Sub-case $A1$: $\lambda_{PD} \geq \lambda_{PN}$ and $\lambda_{DD} \geq \lambda_{NN}$. In this sub-case, the original condition $\lambda_D \geq \lambda_N$ is automatically satisfied. For this sub-case, it is clear that the states PN and NP can be fully alternated along with the PD and DP using scheme $S_2^{3/2}$ to achieve $\frac{3}{2}$ s.d.o.f. The remaining fraction of time for PD (and DP) is hence: $\lambda_{PD}-\lambda_{PN}$. The state NN can be fully utilized along with DD to achieve 1 s.d.o.f. using the scheme S_2^1. The DN and ND states are alternated with each other to achieve 1 s.d.o.f. Thus, we achieve the following sum s.d.o.f.:

$$d_1+d_2 = \underbrace{2}_{S^2}\times\lambda_{PP}+\underbrace{\frac{3}{2}}_{S_2^{3/2}}\times(2\lambda_{PD}+2\lambda_{PN})+\underbrace{1}_{S_2^1}\times(\lambda_{DD}+\lambda_{NN})+2\lambda_{DN}$$

$$= 2\lambda_{PP}+3\lambda_{PD}+3\lambda_{PN}+\lambda_{DD}+\lambda_{NN}+2\lambda_{DN} \tag{197}$$

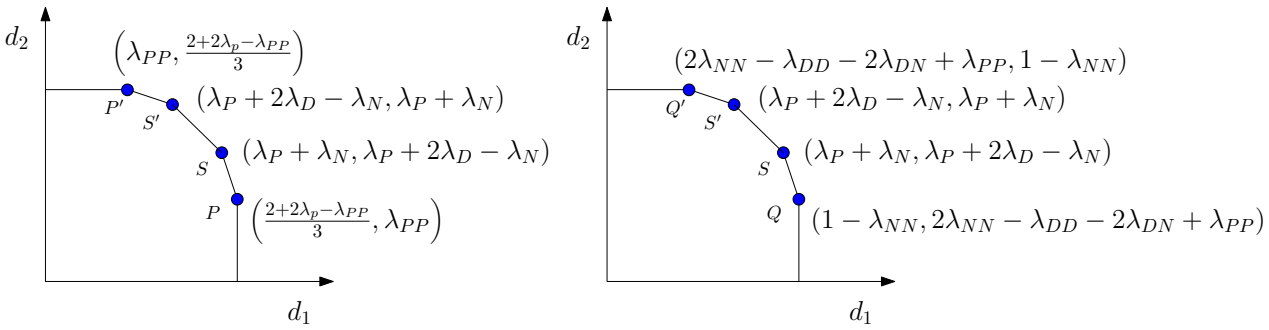
$$= 1+\lambda_P. \tag{198}$$

Sub-case $A2$: $\lambda_{PD} \geq \lambda_{PN}$, $\lambda_{DD} \leq \lambda_{NN}$. As in sub-case $A1$, we can fully alternate the PN and NP states with the PD and DP states using the scheme $S_2^{3/2}$ to achieve the s.d.o.f. of $\frac{3}{2}$. Since $\lambda_{DD} \leq \lambda_{NN}$, we instead fully alternate the state DD along with NN using scheme S_2^1 to achieve a sum s.d.o.f. of 1. The remaining fraction of the NN state is $\lambda_{NN} - \lambda_{DD}$ which can be alternated with the remaining fraction of (PD, DP), which is $\lambda_{PD} - \lambda_{PN}$ as long as $\lambda_{PD} - \lambda_{PN} \geq \lambda_{NN} - \lambda_{DD}$. This achieves $\frac{4}{3}$ sum s.d.o.f. Indeed, this is feasible as this is precisely the condition $\lambda_D \geq \lambda_N$. The DN and ND states are alternated with each other to achieve 1 s.d.o.f.

$$\begin{aligned} d_1 + d_2 = &\underbrace{2}_{S^2} \times \lambda_{PP} + \underbrace{\frac{3}{2}}_{S_2^{3/2}} \times (4\lambda_{PN}) + \underbrace{1}_{S_2^1} \times (2\lambda_{DD}) + 2\lambda_{DN} \\ &+ \underbrace{\frac{4}{3}}_{S_1^{4/3}} \times (3(\lambda_{NN} - \lambda_{DD})) + \underbrace{\frac{3}{2}}_{S_1^{3/2}} \times 2(\lambda_{PD} - \lambda_{PN} - \lambda_{NN} + \lambda_{DD}) \qquad (199) \\ = &\, 2\lambda_{PP} + 6\lambda_{PN} + 2\lambda_{DD} + 4\lambda_{NN} - 4\lambda_{DD} + 3\lambda_{PD} + 3\lambda_{DD} - 3\lambda_{PN} - 3\lambda_{NN} + 2\lambda_{DN} \\ &\qquad (200) \\ = &\, 2\lambda_{PP} + 3\lambda_{PD} + 3\lambda_{PN} + \lambda_{DD} + \lambda_{NN} + 2\lambda_{DN} \qquad (201) \\ = &\, 1 + \lambda_P. \qquad (202) \end{aligned}$$

Sub-case $A3$: $\lambda_{PD} \leq \lambda_{PN}$, $\lambda_{DD} \geq \lambda_{NN}$. Unlike the previous two sub-cases, here, we cannot fully alternate the PN and NP states with the PD and DP states. Instead, we fully use up the PD and DP states with a part of the PN and NP states using scheme $S_2^{3/2}$ to achieve the sum s.d.o.f. of $\frac{3}{2}$. The remaining duration of PN (or the NP) state is $\lambda_{PN} - \lambda_{PD}$. Now, we can also fully alternate the NN state with DD since $\lambda_{DD} \geq \lambda_{NN}$ using the scheme S_2^1 to achieve the sum s.d.o.f. of 1; and thus, the remaining fraction of DD state is $\lambda_{DD} - \lambda_{NN}$. We now alternate the remaining PN and NP states with the remaining DD state using the scheme $S_2^{4/3}$ to achieve the sum s.d.o.f. of $\frac{4}{3}$. For this to be feasible, we require $\lambda_{DD} - \lambda_{NN} \geq \lambda_{PN} - \lambda_{PD}$ which is again precisely the condition $\lambda_D \geq \lambda_N$. The remaining DD state achieves sum s.d.o.f. of 1 using scheme S_1^1. The DN and ND states are alternated with each other to achieve 1 s.d.o.f.

$$\begin{aligned} d_1 + d_2 = &\underbrace{2}_{S^2} \times \lambda_{PP} + \underbrace{\frac{3}{2}}_{S_2^{3/2}} \times (4\lambda_{PD}) + \underbrace{1}_{S_2^1} \times (2\lambda_{NN}) \\ &+ \underbrace{\frac{4}{3}}_{S_2^{4/3}} \times (3(\lambda_{PN} - \lambda_{PD})) + \underbrace{1}_{S_1^1} \times (\lambda_{DD} - \lambda_{NN} - \lambda_{PN} + \lambda_{PD}) + 2\lambda_{DN} \qquad (203) \end{aligned}$$



(a) $\lambda_{DD}+2\lambda_{DN} \geq 2\lambda_{NN}$. (b) $\lambda_{DD}+2\lambda_{DN} \leq 2\lambda_{NN}$.

Figure 11: s.d.o.f. regions in case B when $3d_1+d_2$ and d_1+3d_2 bounds are partially active.

$$= 2\lambda_{PP}+6\lambda_{PD}+2\lambda_{NN}+4\lambda_{PN}-4\lambda_{PD}+\lambda_{PD}+\lambda_{DD}-\lambda_{PN}-\lambda_{NN}+2\lambda_{DN} \tag{204}$$

$$= 2\lambda_{PP}+3\lambda_{PD}+3\lambda_{PN}+\lambda_{DD}+\lambda_{NN}+2\lambda_{DN} \tag{205}$$

$$= 1+\lambda_P. \tag{206}$$

Hence, for Case A, i.e., when $\lambda_D \geq \lambda_N$, we have the complete characterization of the s.d.o.f. region.

5.2 Achievability for Case B: $\lambda_N > \lambda_D$

In this case, the $3d_1+d_2/d_1+3d_2$ bounds are inactive at the symmetric sum rate point. However, these $3d_1+d_2/d_1+3d_2$ bounds play a role at other points in the region, in particular, when one of the users requires full secure rate, the $3d_1+d_2/d_1+3d_2$ bounds are relevant in some cases. Thus, these bounds are still partially relevant. Based on whether the $3d_1+d_2/d_1+3d_2$ bounds are partially relevant or completely irrelevant, we divide our achievability into two broad cases:

1. $3d_1+d_2$ bounds are partially relevant, at the point where one user requires full secret rate,

2. $3d_1+d_2$ bounds are completely irrelevant to the region.

Now let us investigate each of these two cases individually.

5.2.1 When $3d_1+d_2$ Bounds are Partially Relevant

This case happens when the intersection of the lines defined by the $3d_1+d_2$ bound and the single user bound is inside the region defined by the lines $d_1=0$, $d_2=0$, single user bounds and the d_1+d_2 bound. We note that this depends on which of the single user bounds is active, giving rise to two cases, as shown in Fig. 11:

- $1-\lambda_{NN} \geq \frac{2+2\lambda_P-\lambda_{PP}}{3}$, in which case, the $3d_1+d_2$ bounds are always relevant, since $\lambda_{PP} \leq 2(\lambda_P+\lambda_D)-\frac{2+2\lambda_P-\lambda_{PP}}{3}$. In this case, when one user requires full rate, it suffices to achieve extremal point given by:

$$P:(d_1,d_2)=\left(\frac{2+2\lambda_P-\lambda_{PP}}{3},\lambda_{PP}\right), \tag{207}$$

- $1-\lambda_{NN} \leq \frac{2+2\lambda_P-\lambda_{PP}}{3}$, in which case, the $3d_1+d_2$ bounds are relevant as long as $\lambda_{NN} \leq \lambda_D$. We will need to show the achievability of one of the extremal points when one of the users requires full rate, given by:

$$Q:(d_1,d_2)=(1-\lambda_{NN},\lambda_{PP}+(2\lambda_{NN}-2\lambda_{DN}-\lambda_{DD})). \tag{208}$$

However, we note that in both cases, the extremal points that achieve the sum rate are defined by the intersection of the lines $3d_1+d_2=2+2\lambda_P$ and $d_1+d_2=2(\lambda_P+\lambda_D)$. These points are symmetric with respect to the line $d_1=d_2$ and it suffices to show the achievability of either one of them. As shown in the figures, it suffices to achieve the point

$$S:(d_1,d_2)=(\lambda_P+\lambda_N,\lambda_P+2\lambda_D-\lambda_N). \tag{209}$$

Thus, to show the achievability of the full region, we need to show how the points P, Q and S are achieved in their relevant cases. We will begin with point S since it remains unaffected by which of the single user bounds is active.

The sum rate point S:

Now we are effectively operating under the constraint $\lambda_{NN} \leq \lambda_D \leq \lambda_N$, and wish to achieve the point $(\lambda_P+\lambda_N,\lambda_P+2\lambda_D-\lambda_N)$. From this condition it is not immediately clear how the constituent schemes should be jointly utilized. Hence we focus on the second half of the inequality, which simplifies to $\lambda_{PD}+\lambda_{DD} \leq \lambda_{PN}+\lambda_{NN}$, and break this condition into three mutually exclusive cases:

- Sub-case $B1$: $\lambda_{PD} \leq \lambda_{PN}$ and $\lambda_{DD} \leq \lambda_{NN}$,
- Sub-case $B2$: $\lambda_{PD} \geq \lambda_{PN}$ and $\lambda_{DD} \leq \lambda_{NN}$,
- Sub-case $B3$: $\lambda_{PD} \leq \lambda_{PN}$ and $\lambda_{DD} \geq \lambda_{NN}$.

Now let us consider each case one by one:

Sub-case $B1$: $\lambda_{PD} \leq \lambda_{PN}$ and $\lambda_{DD} \leq \lambda_{NN}$: In this case, the full DD state will be used up with a part of the NN state using scheme S_2^1 to achieve the rate pair $\left(\frac{1}{2},\frac{1}{2}\right)$. The duration of the remaining NN state is $(\lambda_{NN}-\lambda_{DD})$. Now if $\lambda_{NN}-\lambda_{DD} \leq \lambda_{DN}$, this remaining NN state can be fully used up with the DN and ND states using scheme $S_3^{2/3}$ achieving the pair $(\frac{2}{3},0)$.

The remaining DN and ND states achieve the pair $(\frac{1}{2}, \frac{1}{2})$ using the scheme S_3^1. The PD and DP states are fully alternated with the PN and NP states using scheme $S_2^{3/2}$ to achieve the pair $\left(\frac{3}{4}, \frac{3}{4}\right)$. The remaining PN and NP states achieve the pair $(1, 0)$. The rate pair achieved then is

$$\begin{aligned}
d_1 =&\lambda_{PP} + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times 4\lambda_{PD} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + 1 \times 2(\lambda_{PN} - \lambda_{PD}) + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3(\lambda_{NN} - \lambda_{DD}) \\
&+ \underbrace{\frac{1}{2}}_{S_3^1} \times 2(\lambda_{DN} - \lambda_{NN} + \lambda_{DD}) \\
=&\lambda_{PP} + \lambda_{PD} + \lambda_{DN} + \lambda_{NN} + 2\lambda_{PN} \\
=&\lambda_P + \lambda_N \qquad (210) \\
d_2 =&\lambda_{PP} + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times 4\lambda_{PD} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + \underbrace{\frac{1}{2}}_{S_3^1} \times 2(\lambda_{DN} - \lambda_{NN} + \lambda_{DD}) \\
=&\lambda_{PP} + 3\lambda_{PD} + 2\lambda_{DD} + \lambda_{DN} - \lambda_{NN} \\
=&\lambda_P + 2\lambda_D - \lambda_N. \qquad (211)
\end{aligned}$$

If on the other hand, $\lambda_{NN} - \lambda_{DD} \geq \lambda_{DN}$, the remaining state NN cannot be fully alternated with the states DN and ND. However, $\lambda_{NN} \leq \lambda_{DN} + \lambda_{DD} + \lambda_{PD}$ from our original condition. Therefore, the full DN and ND states are alternated with a part of the NN state using scheme $S_3^{2/3}$ achieving the pair $(\frac{2}{3}, 0)$. The remaining duration of the NN state is $(\lambda_{NN} - \lambda_{DD} - \lambda_{DN})$, which can be fully alternated with the PD and DP states using the scheme $S_1^{4/3}$ achieving the pair $(\frac{2}{3}, \frac{2}{3})$. The remaining PD and DP states can be alternated with the PN and NP states using scheme $S_2^{3/2}$ achieving the point $(\frac{3}{4}, \frac{3}{4})$. The rest of the PN and NP states achieve the point $(1, 0)$. Thus, we have,

$$\begin{aligned}
d_1 =&\lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3\lambda_{DN} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3(\lambda_{NN} - \lambda_{DN} - \lambda_{DD}) \\
&+ \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times 4(\lambda_{PD} - (\lambda_{NN} - \lambda_{DN} - \lambda_{DD})) + 1 \times 2(\lambda_{PN} - \lambda_{PD} + (\lambda_{NN} - \lambda_{DN} - \lambda_{DD})) \\
=&\lambda_P + \lambda_N \qquad (212) \\
d_2 =&\lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3(\lambda_{NN} - \lambda_{DN} - \lambda_{DD}) + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times 4(\lambda_{PD} - (\lambda_{NN} - \lambda_{DN} - \lambda_{DD})) \\
=&\lambda_P + 2\lambda_D - \lambda_N. \qquad (213)
\end{aligned}$$

Sub-case $B2$: $\lambda_{PD} \geq \lambda_{PN}$ and $\lambda_{DD} \leq \lambda_{NN}$: In this case, since $\lambda_{NN} \geq \lambda_{DD}$, the entire DD state is alternated with a portion of the NN state using scheme S_2^1 to achieve the s.d.o.f. pair $\left(\frac{1}{2}, \frac{1}{2}\right)$. The remaining duration of the NN state is $\lambda_{NN} - \lambda_{DD}$. Now if $\lambda_{NN} - \lambda_{DD} \leq \lambda_{PD}$, the remaining NN state is used with a part of the PD and DP states in scheme $S_1^{4/3}$ achieving the pair $\left(\frac{2}{3}, \frac{2}{3}\right)$. The remaining portion of the PD and DP states can then be utilized with the PN and NP states using scheme $S_2^{3/2}$ achieving the pair $\left(\frac{3}{4}, \frac{3}{4}\right)$. The remaining PN and NP states are utilized to just achieve the rate pair $(1, 0)$. The DN and ND states are used to achieve the pair $(\frac{1}{2}, \frac{1}{2})$ using the scheme S_3^1. Thus, we have,

$$\begin{aligned} d_1 =& \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times (2\lambda_{DD}) + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times (3(\lambda_{NN} - \lambda_{DD})) + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times (4\,(\lambda_{PD} - (\lambda_{NN} - \lambda_{DD}))) \\ & + 1 \times (2\lambda_{PN} - 2\,(\lambda_{PD} - (\lambda_{NN} - \lambda_{DD}))) + \frac{1}{2} \times 2\lambda_{DN} && (214) \\ =& \lambda_P + \lambda_N && (215) \\ d_2 =& \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times (2\lambda_{DD}) + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times (3(\lambda_{NN} - \lambda_{DD})) + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times (4\,(\lambda_{PD} - (\lambda_{NN} - \lambda_{DD}))) \\ & + \frac{1}{2} \times 2\lambda_{DN} && (216) \\ =& \lambda_{PP} + 2\lambda_{DD} + 3\lambda_{PD} - \lambda_{NN} + \lambda_{DN} && (217) \\ =& \lambda_P + 2\lambda_D - \lambda_N. && (218) \end{aligned}$$

If on the other hand, $\lambda_{NN} - \lambda_{DD} \geq \lambda_{PD}$, the full PD and DP states will be used up with a part of the remaining NN state using scheme $S_1^{4/3}$ achieving the pair $(\frac{2}{3}, \frac{2}{3})$. The remaining duration of the NN state is $\lambda_{NN} - \lambda_{DD} - \lambda_{PD}$, which is less than λ_{DN} from our original condition. Therefore, this remaining NN state can be fully utilized with the DN and ND states using scheme $S_3^{2/3}$ to achieve the pair $(\frac{2}{3}, 0)$. The remaining DN and ND states achieve the pair $(\frac{1}{2}, \frac{1}{2})$, while the PN and NP states achieve the pair $(1, 0)$. Thus, we have,

$$\begin{aligned} d_1 =& \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3\lambda_{PD} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3(\lambda_{NN} - \lambda_{DD} - \lambda_{PD}) \\ & + \underbrace{\frac{1}{2}}_{S_3^1} \times 2(\lambda_{DN} + \lambda_{DD} + \lambda_{PD} - \lambda_{NN}) + 1 \times 2\lambda_{PN} && (219) \\ =& \lambda_{PP} + \lambda_{PD} + 2\lambda_{PN} + \lambda_{DN} + \lambda_{NN} && (220) \\ =& \lambda_P + \lambda_N && (221) \end{aligned}$$

$$d_2 = \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3\lambda_{PD} + \underbrace{\frac{1}{2}}_{S_3^1} \times 2(\lambda_{DN} + \lambda_{DD} + \lambda_{PD} - \lambda_{NN})$$

$$= \lambda_{PP} + 2\lambda_{DD} + 3\lambda_{PD} + \lambda_{DN} - \lambda_{NN} \tag{222}$$

$$= \lambda_P + 2\lambda_D - \lambda_N. \tag{223}$$

Sub-case $B3$: $\lambda_{PD} \leq \lambda_{PN}$ and $\lambda_{DD} \geq \lambda_{NN}$: To achieve the sum rate point, we should alternate the entire PD and DP states with part of the PN and NP states using the scheme $S_2^{3/2}$. Also the entire NN state should be alternated with the DD state using the scheme S_2^1. The remaining DD state can then be fully utilized with a part of the remaining PN and NP states using scheme $S_2^{4/3}$, since, $\lambda_{DD} - \lambda_{NN} \leq \lambda_{PN} - \lambda_{PD}$. The remaining PN and NP states will be exploited to achieve the s.d.o.f. pair $(1, 0)$. The DN and ND states together achieve the pair $(\frac{1}{2}, \frac{1}{2})$. Thus, we have,

$$d_1 = \lambda_{PP} + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times (4\lambda_{PD}) + \underbrace{\frac{1}{2}}_{S_2^1} \times (2\lambda_{NN}) + \underbrace{\frac{2}{3}}_{S_2^{4/3}} \times (3(\lambda_{DD} - \lambda_{NN})) + \frac{1}{2} \times 2\lambda_{DN}$$

$$+ 1 \times (2(\lambda_{PN} - \lambda_{PD}) - 2(\lambda_{DD} - \lambda_{NN})) \tag{224}$$

$$= \lambda_{PP} + \lambda_{PD} + 2\lambda_{PN} + \lambda_{NN} + \lambda_{DN} \tag{225}$$

$$= \lambda_P + \lambda_N \tag{226}$$

$$d_2 = \lambda_{PP} + \underbrace{\frac{3}{4}}_{S_2^{3/2}} \times (4\lambda_{PD}) + \underbrace{\frac{1}{2}}_{S_2^1} \times (2\lambda_{NN}) + \underbrace{\frac{2}{3}}_{S_2^{4/3}} \times \left(3(\lambda_{DD} - \lambda_{NN}) + \frac{1}{2} \times 2\lambda_{DN}\right) \tag{227}$$

$$= \lambda_{PP} + 3\lambda_{PD} + 2\lambda_{DD} - \lambda_{NN} + \lambda_{DN} \tag{228}$$

$$= \lambda_P + 2\lambda_D - \lambda_N. \tag{229}$$

The points P and Q:

- Point P: Recall that we need to achieve the point $P : \left(\frac{2+2\lambda_P - \lambda_{PP}}{3}, \lambda_{PP}\right)$ when $1 - \lambda_{NN} \geq \frac{2+2\lambda_P - \lambda_{PP}}{3}$, a condition that simplifies to $\lambda_{DD} + 2\lambda_{DN} \geq 2\lambda_{NN}$. To achieve this point, using the state PP, we achieve $(1, 1)$, with PD, DP, PN, NP, we achieve the pair $(1, 0)$. For the states $(\mathsf{DD}, \mathsf{NN}) \sim (\frac{2}{3}, \frac{1}{3})$, and $(\mathsf{DN}, \mathsf{ND}, \mathsf{NN}) \sim (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, we achieve the pair $(\frac{2}{3}, 0)$ by using the schemes $S_2^{2/3}$ and $S_3^{2/3}$, respectively. Essentially, the NN state is used up with the DD state and the DN and ND states to achieve $\frac{2}{3}$ s.d.o.f. for user 1.

  Time sharing yields the following s.d.o.f. pair:

$$d_2 = \lambda_{PP} \tag{230}$$

$$d_1 = \lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_2^{2/3}} (\lambda_{DD} + 2\lambda_{DN} + \lambda_{NN}) \tag{231}$$

$$= 2\lambda_P - \lambda_{PP} + \frac{2}{3}(\lambda_{DD} + 2\lambda_{DN} + \lambda_{NN}) \tag{232}$$

$$= 2\lambda_P - \lambda_{PP} + \frac{2}{3}(1 - 2\lambda_P + \lambda_{PP}) \tag{233}$$

$$= \frac{2 + 2\lambda_P - \lambda_{PP}}{3}. \tag{234}$$

- Point Q: We need to achieve the point $Q : (1 - \lambda_{NN}, \lambda_{PP} + (2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}))$ when $1 - \lambda_{NN} \leq \frac{2+2\lambda_P - \lambda_{PP}}{3}$, or equivalently, when $\lambda_{DD} + 2\lambda_{DN} \leq \lambda_{NN}$ and under the added constraint $\lambda_{NN} \leq \lambda_D$. Here, we consider two further subcases:

  - $\lambda_{NN} \leq \lambda_{DD} + \lambda_{DN}$: In this case, to achieve the point Q, we first use up the full DN and ND states with a part of the NN state using scheme $S_3^{2/3}$. We alternate the remaining $(\lambda_{NN} - \lambda_{DN})$ duration of NN state with the DD state using two schemes: $S_2^{2/3}$ and S_2^1. Note that in this case, $0 \leq 2(\lambda_{DD} + \lambda_{DN} - \lambda_{NN}) \leq \lambda_{DD}$. We use the state DD for duration $2(\lambda_{DD} + \lambda_{DN} - \lambda_{NN})$ and state NN for duration $(\lambda_{DD} + \lambda_{DN} - \lambda_{NN})$ together using scheme $S_2^{2/3}$ to achieve the s.d.o.f. pair $\left(\frac{2}{3}, 0\right)$. The remaining $(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD})$ duration of the state NN is alternated with the remaining $(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD})$ duration of state DD using the scheme S_2^1 to achieve the s.d.o.f. pair $\left(\frac{1}{2}, \frac{1}{2}\right)$. The state PP allows us to achieve the s.d.o.f. pair $(1, 1)$ while the remaining states PD, DP, PN, and NP each achieves $(1, 0)$. Thus, by using time sharing, the s.d.o.f. pair is:

$$d_1 = \lambda_{PP} + 1 \times 2\lambda_{PD} + 1 \times 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3\lambda_{DN} + \underbrace{\frac{2}{3}}_{S_2^{2/3}} \times 3(\lambda_{DD} + \lambda_{DN} - \lambda_{NN}) \tag{235}$$

$$+ \underbrace{\frac{1}{2}}_{S_2^1} \times 2(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}) \tag{236}$$

$$= 1 - \lambda_{NN} \tag{237}$$

$$d_2 = \lambda_{PP} + \underbrace{\frac{1}{2}}_{S_2^1} \times 2(2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD})$$

$$= \lambda_{PP} + (2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}), \tag{238}$$

    which is precisely the point Q.

  - $\lambda_{NN} \geq \lambda_{DD} + \lambda_{DN}$: In this case, the state NN cannot be completely used with the

states DD, DN and ND. But we note that $\lambda_D \geq \lambda_{NN}$. We first use up the DN and ND states by alternating with the NN state using scheme $S_3^{2/3}$. A portion λ_{DD} of the remaining $(\lambda_{NN} - \lambda_{DN})$ duration of the NN state uses up the DD state in scheme S_2^1 achieving the pair $\left(\frac{1}{2}, \frac{1}{2}\right)$. The remaining $(\lambda_{NN} - \lambda_{DN} - \lambda_{DD})$ portion of the NN state is used with the PD and DP states through the scheme $S_1^{4/3}$ to achieve the pair $\left(\frac{2}{3}, \frac{2}{3}\right)$. For the remainder of the state PD, DP and the states PN, NP, we can achieve the pair $(1, 0)$, while $(1, 1)$ is achieved in the PP state. By time sharing, we get

$$\begin{aligned} d_1 =& \lambda_{PP} + 2\lambda_{PN} + \underbrace{\frac{2}{3}}_{S_3^{2/3}} \times 3\lambda_{DN} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3(\lambda_{NN} - \lambda_{DN} - \lambda_{DD}) \\ &+ 2(\lambda_{PD} - \lambda_{NN} + \lambda_{DN} + \lambda_{DD}) + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} \qquad (239) \\ =& 1 - \lambda_{NN} \qquad (240) \\ d_2 =& \lambda_{PP} + \underbrace{\frac{2}{3}}_{S_1^{4/3}} \times 3(\lambda_{NN} - \lambda_{DN} - \lambda_{DD}) + \underbrace{\frac{1}{2}}_{S_2^1} \times 2\lambda_{DD} \qquad (241) \\ =& \lambda_{PP} + 2\lambda_{NN} - 2\lambda_{DN} - \lambda_{DD}, \qquad (242) \end{aligned}$$

which is again the point Q.

Thus, we have achieved the point Q as well.

This completes the achievability of the full region when the $3d_1 + d_2$ bounds are relevant.

5.2.2 When $3d_1 + d_2$ Bounds are Irrelevant

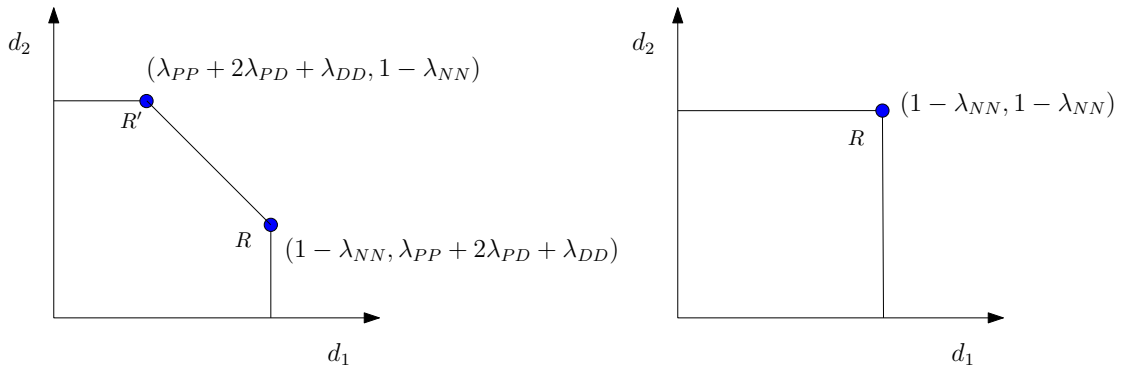
This case occurs when $\lambda_{NN} \geq \lambda_D$. In this case, the single user bounds are

$$\begin{aligned} d_1 \leq& 1 - \lambda_{NN} \qquad (243) \\ d_2 \leq& 1 - \lambda_{NN}, \qquad (244) \end{aligned}$$

and as shown in Fig. 12a the only point to achieve is given by:

$$R : (d_1, d_2) = (1 - \lambda_{NN}, \lambda_{PP} + 2\lambda_{PD} + \lambda_{DD}). \qquad (245)$$

Note that $\lambda_{PP} + 2\lambda_{PD} + \lambda_{DD} \leq 1 - \lambda_{NN}$ with equality if and only if $\lambda_{PN} = \lambda_{DN} = 0$. Thus, it suffices to achieve the point R which goes to the degenerate point $(1 - \lambda_{NN}, 1 - \lambda_{NN})$ when $\lambda_{PN} = \lambda_{DN} = 0$, as shown in Fig. 12b.



(a) $\lambda_{DN}+\lambda_{PN}\neq 0$. (b) $\lambda_{DN}=\lambda_{PN}=0$.

Figure 12: s.d.o.f. regions in case B, when $3d_1+d_2$ and d_1+3d_2 bounds are completely irrelevant.

To achieve this point, we alternate part of the NN state with the DD state using scheme S_2^1 to achieve the pair $\left(\frac{1}{2},\frac{1}{2}\right)$, and with the PD and DP states using the scheme $S_1^{4/3}$ to achieve the pair $\left(\frac{2}{3},\frac{2}{3}\right)$ and with the DN and ND states using the scheme $S_3^{2/3}$ to achieve the pair $(\frac{2}{3},0)$. The remaining NN state is left unused. The PN and NP states, if available, is used to achieve the s.d.o.f. pair $(1,0)$. Thus, we have,

$$\begin{aligned} d_1 =& \lambda_{PP}+\underbrace{\frac{1}{2}}_{S_2^1}\times(2\lambda_{DD})+\underbrace{\frac{2}{3}}_{S_1^{4/3}}\times(3\lambda_{PD})+\underbrace{\frac{2}{3}}_{S_3^{2/3}}\times 3\lambda_{DN}+1\times 2\lambda_{PN} \\ =& 1-\lambda_{NN} \qquad (246) \\ d_2 =& \lambda_{PP}+\underbrace{\frac{1}{2}}_{S_2^1}\times(2\lambda_{DD})+\underbrace{\frac{2}{3}}_{S_1^{4/3}}\times(3\lambda_{PD}) \qquad (247) \\ =& \lambda_{PP}+\lambda_{DD}+2\lambda_{PD} \qquad (248) \\ =& 1-\lambda_{NN} \text{ if } \lambda_{PN}=\lambda_{DN}=0. \qquad (249) \end{aligned}$$

This completes the proof of the achievability.

6 Proof of the Converse

6.1 Local Statistical Equivalence Property and Associated Lemma

We introduce a property of the channel which we call *local statistical equivalence*. Let us focus on the channel output of receiver 2 corresponding to the state PD and DD at time t:

$$Z_{pd}(t)=\mathbf{H}_{2,pd}(t)\mathbf{X}_{pd}(t)+N_{2,pd}(t) \qquad (250)$$

$$Z_{dd}(t) = \mathbf{H}_{2,dd}(t)\mathbf{X}_{dd}(t) + N_{2,dd}(t). \tag{251}$$

Now consider $(\tilde{\mathbf{H}}_{2,pd}(t), \tilde{\mathbf{H}}_{2,dd}(t))$, $(\tilde{N}_{2,pd}(t), \tilde{N}_{2,dd}(t))$, which are i.i.d. as $(\mathbf{H}_{2,pd}(t), \mathbf{H}_{2,dd}(t))$ and $(N_{2,pd}(t), N_{2,dd}(t))$, respectively. Using these random variables, we define artificial channel outputs as:

$$\tilde{Z}_{pd}(t) = \tilde{\mathbf{H}}_{2,pd}(t)\mathbf{X}_{pd}(t) + \tilde{N}_{2,pd}(t) \tag{252}$$

$$\tilde{Z}_{dd}(t) = \tilde{\mathbf{H}}_{2,dd}(t)\mathbf{X}_{dd}(t) + \tilde{N}_{2,dd}(t). \tag{253}$$

Let $\Omega = (\mathbf{H}^n, \tilde{\mathbf{H}}^n)$. Now the *local statistical equivalence* property is the following:

$$h(Z_{pd}(t), Z_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega) = h(\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega). \tag{254}$$

This property shows that if we consider the outputs of a receiver for such states in which it supplies delayed CSIT, then the entropy of the channel outputs conditioned on the past outputs is the same as that of another artificial receiver whose channel is distributed identically as the original receiver. Note that in an alternating CSIT setting, we focus on only the states in which the receiver provides delayed CSIT; hence we call it *local*. The original and artificial receivers have *statistically equivalent* channels in the sense that the conditional differential entropies of the outputs at the real and the artificial receivers given the past outputs are equal. The proof of this property is given in Appendix A. We next present the following lemma which together with the local statistical equivalence property is instrumental in the converse proofs.

Lemma 1 *For our channel model, with CSIT alternating among the states* DD*,* PD *and* DP *we have:*

$$h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|Z^n, \Omega) \tag{255}$$

$$2h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|\Omega) \tag{256}$$

$$h(Y^n|\Omega) \dot{\geq} h(Z_{dp}^n, Z_{dd}^n|Y^n, \Omega) \tag{257}$$

$$2h(Y^n|\Omega) \dot{\geq} h(Z_{dp}^n, Z_{dd}^n|\Omega), \tag{258}$$

where $a \dot{\geq} b$ *denotes* $\lim\limits_{P\to\infty} \frac{a}{\log P} \geq \lim\limits_{P\to\infty} \frac{b}{\log P}$.

This lemma is proved in Appendix B.

In the following sections, we use the local statistical equivalence property along with Lemma 1 to prove the bounds on individual d.o.f. d_1 and d_2, the sum d.o.f. $(d_1 + d_2)$ and the weighted d.o.f. $3d_1 + d_2$ and $d_1 + 3d_2$.

6.2 The Single User Bounds

We recall the single user bounds in (14)-(15):

$$d_1 \leq \min\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}, 1-\lambda_{NN}\right) \tag{259}$$

$$d_2 \leq \min\left(\frac{2+2\lambda_P-\lambda_{PP}}{3}, 1-\lambda_{NN}\right). \tag{260}$$

6.2.1 Proof of $d_i \leq \frac{2+2\lambda_P-\lambda_{PP}}{3}$, $i=1,2$

In this section, we prove the following single-user bounds:

$$d_1 \leq \frac{2+2\lambda_P-\lambda_{PP}}{3} = \frac{2+2\lambda_P+2\lambda_{PD}+2\lambda_{PN}}{3} \tag{261}$$

$$d_2 \leq \frac{2+2\lambda_P-\lambda_{PP}}{3} = \frac{2+2\lambda_P+2\lambda_{PD}+2\lambda_{PN}}{3}. \tag{262}$$

To do so, we enhance the transmitter in the following way:

- First, if in any state, the transmitter has perfect CSIT from any of the users, we provide perfect CSI for the other user too, that is, the states $\mathsf{PP}, \mathsf{PD}, \mathsf{DP}, \mathsf{PN}, \mathsf{NP}$ are all enhanced to the state PP.
- Next, we enhance all the remaining states, (i.e., $\mathsf{DD}, \mathsf{DN}, \mathsf{ND}, \mathsf{NN}$) to DD.

The enhanced channel has two states: PP occurring for $\lambda_{pp} = \lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN}$ (using symmetry of the alternation), and DD occurring for the remaining fraction of the time. Now, we have the following lemma for such a channel with only PP and DD states.

Lemma 2 *Consider the two-user MISO BCCM with only two states:* PP *and* DD *occurring for* λ_{pp} *and* λ_{dd} *fractions of time, respectively, such that* $\lambda_{pp}+\lambda_{dd}=1$. *Then,*

$$d_1 \leq \frac{2+\lambda_{pp}}{3} \tag{263}$$

$$d_2 \leq \frac{2+\lambda_{pp}}{3}. \tag{264}$$

The proof of this lemma is provided in Appendix C.1.

Now using $\lambda_{pp} = \lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN}$ in Lemma 2, we get the bounds in (261)-(262).

6.2.2 Proof of $d_i \leq 1-\lambda_{NN}$, $i=1,2$

In this section, we prove the following single user bounds:

$$d_1 \leq 1-\lambda_{NN} \tag{265}$$

$$d_2 \leq 1 - \lambda_{NN}. \tag{266}$$

To prove these, we again enhance the transmitter, but in a different way. We provide the transmitter with perfect CSIT in every state except the NN state, that is, every state except the NN state is enhanced to the PP state. Thus, we end up with a system with two states: PP occurring for $1-\lambda_{NN}$ fraction of the time and NN occurring for λ_{NN} fraction of the time. Note that since there is no delayed CSIT in the enhanced system, there is no feedback. For such a system we have the following lemma.

Lemma 3 *For the two-user MISO BCCM with only two states:* PP *and* NN *occurring for* $1-\lambda_{nn}$ *and* λ_{nn} *fractions of time, respectively, and no feedback,*

$$d_1 \leq 1 - \lambda_{nn} \tag{267}$$

$$d_2 \leq 1 - \lambda_{nn}. \tag{268}$$

The proof of this lemma is provided in Appendix C.2.

Using $\lambda_{nn} = \lambda_{NN}$ in Lemma 3, we get the bounds in (265)-(266).

Combining the bounds in (261)-(262) and (265)-(266), we have the bounds in (14)-(15).

6.3 Proof of $d_1 + d_2$ Bound

Recall the sum s.d.o.f. bound from (18):

$$d_1 + d_2 \leq 2(\lambda_P + \lambda_D). \tag{269}$$

The original system model has nine possible states, namely, PP, DD, NN, DP, PD, PN, NP, DN, and ND. We enhance the transmitter in the following way: whenever in any state, the transmitter receives delayed CSI of a channel, we provide perfect CSI of the channel to the transmitter; in other words, we convert each D state to a P state. This clearly does not decrease the secrecy capacity (and thus, the s.d.o.f. region). Also note that the enhanced system does not have any delayed CSIT, and hence no feedback. Now the enhanced system has only four states: PP, PN, NP, NN, occurring for $\lambda_{pp} = \lambda_{PP} + \lambda_{DD} + \lambda_{DP} + \lambda_{PD}$, $\lambda_{pn} = \lambda_{PN} + \lambda_{DN}$, $\lambda_{np} = \lambda_{NP} + \lambda_{ND}$ and $\lambda_{nn} = \lambda_{NN}$ fractions of time, respectively. For such a system with four states we have the following lemma:

Lemma 4 *Consider the two-user MISO BCCM with only four of the nine states:* PP*,* PN*,* NP *and* NN *occurring for* λ_{pp}*,* λ_{pn}*,* λ_{np} *and* λ_{nn} *fractions of the time, with* $\lambda_{pp} + \lambda_{pn} + \lambda_{np} + \lambda_{nn} = 1$*. Also, assume there is no feedback. Then,*

$$d_1 + d_2 \leq 2\lambda_{pp} + \lambda_{pn} + \lambda_{np}. \tag{270}$$

Proof of this lemma is presented in Appendix C.3.

Thus, using $\lambda_{pp} = \lambda_{PP} + \lambda_{DD} + \lambda_{DP} + \lambda_{PD}$, $\lambda_{pn} = \lambda_{PN} + \lambda_{DN}$, $\lambda_{np} = \lambda_{NP} + \lambda_{ND}$ and $\lambda_{nn} = \lambda_{NN}$ in Lemma 4, we have,

$$d_1 + d_2 \leq 2(\lambda_{PP} + \lambda_{DP} + \lambda_{PD} + \lambda_{DD}) + \lambda_{PN} + \lambda_{DN} + \lambda_{NP} + \lambda_{ND} \tag{271}$$

$$= 2(\lambda_P + \lambda_D), \tag{272}$$

where (272) follows due to the assumed symmetry: $\lambda_{PD} = \lambda_{DP}$, and this completes the proof of the bound on $d_1 + d_2$.

6.4 Proof of $3d_1 + d_2$ and $d_1 + 3d_2$ Bounds

In this section, we prove the following bounds from (16)-(17):

$$3d_1 + d_2 \leq 2 + 2\lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN} \tag{273}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN}. \tag{274}$$

To do so, we enhance the system in the following way: Whenever in any state, the transmitter has no CSIT from a user, we provide the transmitter delayed CSIT of that user's channel; in other words, we enhance each N state to a D state. After this enhancement, we are left with only four states, namely PP, PD, DP and DD occurring for $\lambda_{pp} = \lambda_{PP}$, $\lambda_{pd} = \lambda_{PD} + \lambda_{PN}$, $\lambda_{dp} = \lambda_{DP} + \lambda_{NP}$ and $\lambda_{dd} = \lambda_{DD} + \lambda_{DN} + \lambda_{ND} + \lambda_{NN}$ fractions of the time, respectively. We have the following lemma for such a system with four states:

Lemma 5 *Consider the two-user MISO BCCM with only four of the nine states:* PP, PD, DP *and* DD *occurring for λ_{pp}, λ_{pd}, λ_{dp} and λ_{dd} fractions of the time, with $\lambda_{pd} = \lambda_{dp}$ and $\lambda_{pp} + \lambda_{pd} + \lambda_{dp} + \lambda_{dd} = 1$. Then,*

$$3d_1 + d_2 \leq 2 + 2\lambda_{pp} + 2\lambda_{pd} \tag{275}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_{pp} + 2\lambda_{pd}. \tag{276}$$

We provide a proof for this lemma in Appendix C.4.

Using $\lambda_{pp} = \lambda_{PP}$, $\lambda_{pd} = \lambda_{PD} + \lambda_{PN}$, $\lambda_{dp} = \lambda_{DP} + \lambda_{NP}$ and $\lambda_{dd} = \lambda_{DD} + \lambda_{DN} + \lambda_{ND} + \lambda_{NN}$ in Lemma 5, and symmetry of the alternating states, we have,

$$3d_1 + d_2 \leq 2 + 2\lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN} \tag{277}$$

$$= 2 + 2\lambda_P \tag{278}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_{PP} + 2\lambda_{PD} + 2\lambda_{PN} \tag{279}$$

$$= 2 + 2\lambda_P, \tag{280}$$

which completes the proofs for the bounds on $3d_1 + d_2$ and $d_1 + 3d_2$.

7 Conclusions

In this paper, we studied the two-user MISO broadcast channel with confidential messages (BCCM) and characterized its secure degrees of freedom (s.d.o.f.) region with alternating channel state information at the transmitter (CSIT). The converse proofs for the s.d.o.f. region presented in the paper are based on novel arguments such as local statistical equivalence property and enhancing the system model in different ways, where each carefully chosen enhancement strictly improves the quality of CSIT in a certain manner. For each such enhanced system, we invoke the local statistical equivalence property and incorporate the confidentiality constraints and obtain corresponding upper bounds on the individual (d_1, d_2), sum $(d_1 + d_2)$ and weighted $(3d_1 + d_2, d_1 + 3d_2)$ s.d.o.f.

To establish the achievability of the s.d.o.f. region, several constituent schemes are developed, where each scheme by itself only operates over a subset of 9 states. The achievability of the optimal s.d.o.f. region is then established by time-sharing between the core constituent schemes. The core constituent schemes not only serve the purpose of establishing the s.d.o.f. region but also highlight the synergies across multiple CSIT states which can be exploited to achieve higher s.d.o.f. in comparison to their individually optimal s.d.o.f. values. Besides highlighting the synergistic benefits of alternating CSIT for secrecy, the optimal s.d.o.f. region also quantifies the information theoretic minimal CSIT required from each user to attain a certain s.d.o.f. value. In addition, we also quantify the loss in d.o.f., as a function of the overall CSIT quality, which must be incurred for incorporating confidentiality constraints.

Appendix A Proof of Local Statistical Equivalence

In this section, we prove the *local statistical equivalence* property:

$$h(Z_{pd}(t), Z_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega) = h(\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega). \tag{281}$$

To this end, first denote the common distribution of $(\mathbf{H}_{2,pd}(t), \mathbf{H}_{2,dd}(t))$, $(\tilde{\mathbf{H}}_{2,pd}(t), \tilde{\mathbf{H}}_{2,dd}(t))$ by F. Let $\Omega = \left\{\mathbf{H}_1(t), \mathbf{H}_2(t), \tilde{\mathbf{H}}_1(t), \tilde{\mathbf{H}}_2(t), t = 1, \ldots, n\right\}$ be the set of all channel vectors upto and including time n. Also, let $\Omega_t = \Omega \backslash \left\{\mathbf{H}_{2,pd}(t), \tilde{\mathbf{H}}_{2,pd}(t), \mathbf{H}_{2,dd}(t), \tilde{\mathbf{H}}_{2,dd}(t)\right\}$. We have,

$$h(Z_{pd}(t), Z_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}\Omega)$$

$$=\mathbb{E}_F\left[h(Z_{pd}(t), Z_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega_t, \tilde{\mathbf{H}}_{2,pd}(t), \tilde{\mathbf{H}}_{2,dd}(t), \mathbf{H}_{2,pd}(t)=\mathbf{h}(t), \mathbf{H}_{2,dd}(t)=\mathbf{g}(t))\right] \tag{282}$$

$$=\mathbb{E}_F\left[h(\mathbf{h}(t)\mathbf{X}_{pd}(t)+N_{2,pd}(t), \mathbf{g}(t)\mathbf{X}_{dd}(t)+N_{2,dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega_t)\right] \tag{283}$$

$$=\mathbb{E}_F\left[h(\mathbf{h}(t)\mathbf{X}_{pd}(t)+\tilde{N}_{2,pd}(t), \mathbf{g}(t)\mathbf{X}_{dd}(t)+\tilde{N}_{2,dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega_t)\right] \tag{284}$$

$$=\mathbb{E}_F\Big[h(\mathbf{h}(t)\mathbf{X}_{pd}(t)+\tilde{N}_{2,pd}(t), \mathbf{g}(t)\mathbf{X}_{dd}(t)+\tilde{N}_{2,dd}(t)|Z_{pd}^{t-1}, Z_{pd}^{t-1}, \Omega_t, \tilde{\mathbf{H}}_{2,pd}(t)=\mathbf{h}(t), \tilde{\mathbf{H}}_{2,dd}(t)=\mathbf{g}(t))\Big] \tag{285}$$

$$=\mathbb{E}_F\left[h(\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega_t, \mathbf{H}_{2,pd}(t), \mathbf{H}_{2,dd}(t), \tilde{\mathbf{H}}_{2,pd}(t)=\mathbf{h}(t), \tilde{\mathbf{H}}_{2,dd}(t)=\mathbf{g}(t))\right] \tag{286}$$

$$=h(\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z_{pd}^{t-1}, Z_{dd}^{t-1}, \Omega), \tag{287}$$

where (283) follows because $(\mathbf{X}_{pd}(t), \mathbf{X}_{dd}(t))$ does not depend on $\Big(\mathbf{H}_{2,pd}(t), \tilde{\mathbf{H}}_{2,pd}(t), \mathbf{H}_{2,dd}(t), \tilde{\mathbf{H}}_{2,dd}(t)\Big)$, (284) follows since the additive noises $(N_{2,pd}(t), N_{2,dd}(t))$ and $(\tilde{N}_{2,pd}(t), \tilde{N}_{2,dd}(t))$ are i.i.d. and independent of all other random variables, (285)-(286) follow since $(\mathbf{H}_{2,pd}(t), \mathbf{H}_{2,dd}(t))$ and $(\tilde{\mathbf{H}}_{2,pd}(t), \tilde{\mathbf{H}}_{2,dd}(t))$ have the same distribution F and the fact that $(\mathbf{X}_{pd}(t), \mathbf{X}_{dd}(t))$ does not depend on $(\mathbf{H}_{2,pd}(t), \tilde{\mathbf{H}}_{2,pd}(t), \mathbf{H}_{2,pd}(t), \tilde{\mathbf{H}}_{2,dd}(t))$.

Appendix B Proof of Lemma 1

We consider the scenario in which there are only three CSIT states, namely DD, PD and DP. For such a specific alternating CSIT model, we define the channel outputs as:

$$Z^n \triangleq \left(Z_{dd}^n, Z_{pd}^n, Z_{dp}^n\right)$$
$$Y^n \triangleq \left(Y_{dd}^n, Y_{pd}^n, Y_{dp}^n\right).$$

Also let Ω denote the set of all channel vectors upto and including time n, that is, in other words, $\Omega = \Big\{\mathbf{H}_1(t), \mathbf{H}_2(t), \tilde{\mathbf{H}}_1(t), \tilde{\mathbf{H}}_2(t), t=1,\ldots,n\Big\}$. We wish to prove that with CSIT alternating among the states DD, PD and DP we have:

$$h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|Z^n, \Omega) \tag{288}$$
$$2h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|\Omega) \tag{289}$$
$$h(Y^n|\Omega) \dot{\geq} h(Z_{dp}^n, Z_{dd}^n|Y^n, \Omega) \tag{290}$$
$$2h(Y^n|\Omega) \dot{\geq} h(Z_{dp}^n, Z_{dd}^n|\Omega). \tag{291}$$

First we note that due to symmetry, it suffices to prove (288) and (289). We proceed as follows:

$$h(Z^n|\Omega) = h(Z^n_{pd}, Z^n_{dd}|\Omega) + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}\Omega) \tag{292}$$

$$= \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega). \tag{293}$$

Using the *local statistical equivalence* property, we get,

$$h(Z^n|\Omega) = \sum_{t=1}^{n} h(\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega). \tag{294}$$

Adding (293) and (294), and lower bounding, we get,

$$\begin{aligned}
2h(Z^n|\Omega) \geq & \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t), \tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) + 2h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}\Omega) \\
\geq & \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t), \tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega) \\
& + no(\log P) \qquad (295) \\
= & \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t), \tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t), Y_{pd}(t), Y_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) \\
& - \sum_{t=1}^{n} h(Y_{pd}(t), Y_{dd}(t)|Z_{pd}(t), Z_{dd}(t), \tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t), Z^{t-1}_{pd}, Z^{t-1}_{dd}\Omega) \\
& + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega) + no(\log P) \qquad (296) \\
\geq & \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t), Y_{pd}(t), Y_{dd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, \Omega) + h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega) \\
& + no(\log P) \qquad (297) \\
\geq & \sum_{t=1}^{n} h(Z_{pd}(t), Z_{dd}(t), Y_{dd}(t)Y_{pd}(t)|Z^{t-1}_{pd}, Z^{t-1}_{dd}, Y^{t-1}_{pd}, Y^{t-1}_{dd}\Omega) \\
& + h(Z^n_{dp}|Z^n_{pd}, Y^n_{pd}, Y^n_{dd}, Z^n_{dd}, \Omega) + no(\log P) \qquad (298) \\
= & h(Z^n_{pd}, Z^n_{dd}, Y^n_{pd}, Y^n_{dd}|\Omega) + h(Z^n_{dp}|Z^n_{pd}, Y^n_{pd}, Z^n_{dd}, Y^n_{dd}\Omega) + no(\log P) \qquad (299) \\
= & h(Z^n, Y^n_{pd}, Y^n_{dd}|\Omega) + no(\log P), \qquad (300)
\end{aligned}$$

where (295) follows by noting that

$$h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \Omega) \geq h(Z^n_{dp}|Z^n_{pd}, Z^n_{dd}, \mathbf{X}^n, \Omega) = no(\log P) \tag{301}$$

and (296) follows since given $(Z_{pd}(t), \tilde{Z}_{pd}(t), Z_{dd}(t), \tilde{Z}_{dd}(t))$, one can reconstruct $(\mathbf{X}_{pd}(t), \mathbf{X}_{dd}(t))$ and hence $(Y_{pd}(t), Y_{dd}(t))$ within noise distortion, implying that

$$h(Y_{pd}(t), Y_{dd}(t)|Z_{pd}(t), Z_{dd}(t)\tilde{Z}_{pd}(t), \tilde{Z}_{dd}(t), Z_{pd}^{t-1}, \Omega) \leq no(\log P). \tag{302}$$

Now both (288) and (289) can be derived from (300). We simply expand the right hand side of (300) in two ways:

$$\begin{aligned} 2h(Z^n|\Omega) \geq & h(Z^n, Y_{pd}^n, Y_{dd}^n|\Omega) + no(\log P) && (303)\\ = & h(Z^n|\Omega) + h(Y_{pd}^n, Y_{dd}^n|Z^n, \Omega) + no(\log P), && (304) \end{aligned}$$

which implies $h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|Z^n, \Omega)$, which is exactly (288). Alternatively from (300), we also have

$$\begin{aligned} 2h(Z^n|\Omega) \geq & h(Y_{pd}^n, Y_{dd}^n|\Omega) + h(Z^n|Y_{pd}^n, Y_{dd}^n\Omega) + no(\log P) && (305)\\ \geq & h(Y_{pd}^n, Y_{dd}^n|\Omega) + no(\log P), && (306) \end{aligned}$$

which implies $2h(Z^n|\Omega) \dot{\geq} h(Y_{pd}^n, Y_{dd}^n|\Omega)$, thus proving the relation in (289). This completes the proof of Lemma 1.

Appendix C Proofs of Lemmas 2-5

C.1 Proof of Lemma 2

Recall that we wish to prove that for the two-user MISO BC with only two states: PP and DD occurring for λ_{pp} and λ_{dd} fractions of time, respectively, such that $\lambda_{pp} + \lambda_{dd} = 1$,

$$d_1 \leq \frac{2 + \lambda_{pp}}{3}, \quad d_2 \leq \frac{2 + \lambda_{pp}}{3}. \tag{307}$$

To do so, we proceed as follows:

$$\begin{aligned} nR_1 \leq & I(W_1; Y_{pp}^n, Y_{dd}^n|\Omega) + no(n) && (308)\\ = & I(W_1; Y_{dd}^n|\Omega) + I(W_1; Y_{pp}^n|Y_{dd}^n, \Omega) + no(n) && (309)\\ \leq & n\lambda_{pp} \log P + I(W_1; Y_{dd}^n|\Omega) + no(n) && (310)\\ \leq & n\lambda_{pp} \log P + I(W_1; Y_{dd}^n, Z_{dd}^n|\Omega) + no(n) && (311)\\ \leq & n\lambda_{pp} \log P + I(W_1; Y_{dd}^n|Z_{dd}^n, \Omega) + no(\log P) + no(n) && (312)\\ \leq & n\lambda_{pp} \log P + h(Y_{dd}^n|Z_{dd}^n, \Omega) + no(\log P) + no(n) && (313)\\ \leq & n\lambda_{pp} \log P + h(Z_{dd}^n|\Omega) + no(\log P) + no(n), && (314) \end{aligned}$$

where (308) follows from decodability of W_1 at receiver 1 and Fano's inequality, (313) follows from confidentiality constraint of message W_1 at receiver 2, and (314) follows from application of Lemma 1.

Starting from (310), we also have

$$\begin{aligned}
nR_1 &\leq n\lambda_{pp}\log P + I(W_1; Y_{dd}^n|\Omega) + no(n) && (315)\\
&\leq n\lambda_{pp}\log P + I(W_1; Y_{dd}^n|\Omega) - I(W_1; Z_{dd}^n|\Omega) + no(\log P) + no(n) && (316)\\
&\leq n\lambda_{pp}\log P + h(Y_{dd}^n|\Omega) - h(Y_{dd}^n|W_1,\Omega) - h(Z_{dd}^n|\Omega) + h(Z_{dd}^n|W_1,\Omega) + no(\log P)\\
&\quad + no(n) && (317)\\
&\leq n\lambda_{pp}\log P + h(Y_{dd}^n|\Omega) - \frac{1}{2}h(Z_{dd}^n|W_1,\Omega) - h(Z_{dd}^n|\Omega) + h(Z_{dd}^n|W_1,\Omega) + no(\log P)\\
&\quad + no(n) && (318)\\
&\leq n\lambda_{pp}\log P + h(Y_{dd}^n|\Omega) + \frac{1}{2}h(Z_{dd}^n|W_1,\Omega) - h(Z_{dd}^n|\Omega) + no(\log P) + no(n) && (319)\\
&\leq n\lambda_{pp}\log P + h(Y_{dd}^n|\Omega) + \frac{1}{2}h(Z_{dd}^n|\Omega) - h(Z_{dd}^n|\Omega) + no(\log P) + no(n) && (320)\\
&= n\lambda_{pp}\log P + h(Y_{dd}^n|\Omega) - \frac{1}{2}h(Z_{dd}^n|\Omega) + no(\log P) + no(n) && (321)\\
&\leq n\lambda_{pp}\log P + n\lambda_{dd}\log P - \frac{1}{2}h(Z_{dd}^n|\Omega) + no(\log P) + no(n), && (322)
\end{aligned}$$

where (316) follows from confidentiality constraint of message W_1 at receiver 2, (318) follows from application of Lemma 1, and (320) follows from the fact that conditioning reduces differential entropy.

Eliminating $h(Z_{dd}^n|\Omega)$ from the bounds (322) and (314), we have,

$$\begin{aligned}
3nR_1 \leq& (3n\lambda_{pp} + 2n\lambda_{dd})\log P + no(\log P) + no(n) && (323)\\
=& (2+\lambda_{pp})n\log P + no(\log P). && (324)
\end{aligned}$$

Now first dividing by n and letting $n \to \infty$, then dividing by $\log P$ and letting $P \to \infty$, we get,

$$d_1 \leq \frac{2+\lambda_{pp}}{3}. \tag{325}$$

By symmetry, we get the same single user bound for user 2, completing the proof of Lemma 2.

C.2 Proof of Lemma 3

We want to show that for the two-user MISO BC with only two states: PP and NN occurring for $1-\lambda_{nn}$ and λ_{nn} fractions of time, respectively,

$$d_1 \leq 1-\lambda_{nn} \tag{326}$$
$$d_2 \leq 1-\lambda_{nn}. \tag{327}$$

To prove this, we note that since there is no feedback, the secrecy capacity depends only on the marginal distributions of channel outputs given the input distribution; [54]. Since the transmitter does not have channel knowledge of any of the users in the state NN, our system with outputs

$$Y^n = (Y_{pp}^n, Y_{nn}^n) \tag{328}$$
$$Z^n = (Z_{pp}^n, Z_{nn}^n) \tag{329}$$

has the same secrecy capacity of a new system with outputs given by

$$Y^n = (Y_{pp}^n, Y_{nn}^n) \tag{330}$$
$$Z^n = (Z_{pp}^n, Y_{nn}^n). \tag{331}$$

Thus, from the secrecy requirement, we get,

$$I(W_1; Y_{nn}^n) = I(W_1; Z_{nn}^n) \leq I(W_1; Z^n) \leq no(\log P). \tag{332}$$

Then we have,

$$\begin{aligned}
nR_1 \leq & I(W_1; Y_{pp}^n, Y_{nn}^n) + no(n) && (333)\\
= & I(W_1; Y_{nn}^n) + I(W_1; Y_{pp}^n | Y_{nn}^n) + no(n) && (334)\\
\leq & I(W_1; Y_{pp}^n | Y_{nn}^n) + no(\log P) + no(n) && (335)\\
\leq & h(Y_{pp}^n | Y_{nn}^n) + no(\log P) + no(n) && (336)\\
\leq & h(Y_{pp}^n) + no(\log P) + no(n) && (337)\\
\leq & n(1-\lambda_{nn}) \log P + no(\log P) + no(n), && (338)
\end{aligned}$$

where, (335) follows from equation (332), (336) follows since $h(Y_{pp}^n | Y_{nn}^n, W_1) \geq h(Y_{pp}^n | Y_{nn}^n, W_1, \mathbf{X}^n) \geq o(\log P)$, and (337) follows since conditioning reduces differential entropy.

Dividing by n, and letting $n \to \infty$, we get,

$$R_1 \leq (1 - \lambda_{nn}) \log P + o(\log P). \tag{339}$$

Dividing by $\log P$ and letting $P \to \infty$, we have,

$$d_1 \leq 1 - \lambda_{nn}. \tag{340}$$

By symmetry, we also have,

$$d_2 \leq 1 - \lambda_{nn}. \tag{341}$$

This completes the proof of Lemma 3.

C.3 Proof of Lemma 4

We wish to prove that for the two-user MISO BC with no feedback and only four of the nine states: PP, PN, NP and NN occurring for λ_{pp}, λ_{pn}, λ_{np} and λ_{nn} fractions of the time, with $\lambda_{pp} + \lambda_{pn} + \lambda_{np} + \lambda_{nn} = 1$,

$$d_1 + d_2 \leq 2\lambda_{pp} + \lambda_{pn} + \lambda_{np}. \tag{342}$$

To that end, for each of the two receivers, we introduce another statistically equivalent receiver. At receiver 1, we introduce a virtual receiver $\tilde{1}$, with channel output denoted by $\tilde{Y}$, while the channel output at the virtual receiver $\tilde{2}$ at receiver 2 is denoted by $\tilde{Z}$. Since the secrecy capacity without feedback depends only on the marginals [54], without loss of generality, we can assume that the channels in the state NN are the same for all receivers. The outputs at each of the receivers are

$$\begin{aligned}
Y^n &= (Y_{pp}^n, Y_{pn}^n, Y_{np}^n, Y_{nn}^n) && (343)\\
Z^n &= (Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n) && (344)\\
\tilde{Y}^n &= (Y_{pp}^n, Y_{pn}^n, \tilde{Y}_{np}^n, Y_{nn}^n) && (345)\\
\tilde{Z}^n &= (Z_{pp}^n, \tilde{Z}_{pn}^n, Z_{np}^n, Y_{nn}^n), && (346)
\end{aligned}$$

where

$$\begin{aligned}
\tilde{Y}_{np}(t) &= \tilde{\mathbf{H}}_{1,np}(t)\mathbf{X}_{np}(t) + \tilde{N}_{1,np}(t) && (347)\\
\tilde{Z}_{pn}(t) &= \tilde{\mathbf{H}}_{2,pn}(t)\mathbf{X}_{pn}(t) + \tilde{N}_{2,pn}(t), && (348)
\end{aligned}$$

such that $\tilde{\mathbf{H}}_{1,np}$, $\tilde{\mathbf{H}}_{2,pn}$ are i.i.d. with the same distribution as $\mathbf{H}_{1,np}$, $\mathbf{H}_{2,pn}$, respectively, and $\tilde{N}_{1,np}$, $\tilde{N}_{2,pn}$ are i.i.d. with same distribution as $N_{1,np}$, $N_{2,pn}$. We upper bound the first receiver's rate as

$$\begin{aligned}
nR_1 \leq & I(W_1; Y_{pp}^n, Y_{pn}^n, Y_{np}^n, Y_{nn}^n|\Omega) + no(n) && (349)\\
= & I(W_1, Y_{pn}^n, Y_{np}^n, Y_{nn}^n|\Omega) + I(W_1, Y_{pp}^n|Y_{pn}^n, Y_{np}^n, Y_{nn}^n, \Omega) && (350)\\
\leq & n\lambda_{pp}\log P + I(W_1, Y_{pn}^n, Y_{np}^n, Y_{nn}^n|\Omega) && (351)\\
= & n\lambda_{pp}\log P + I(W_1; Y_{pn}^n Y_{nn}^n|\Omega) + I(W_1; Y_{np}^n|Y_{pn}^n Y_{nn}^n, \Omega) + no(n) && (352)\\
= & n\lambda_{pp}\log P + I(W_1; Y_{pn}^n, Y_{nn}^n|\Omega) + h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, \Omega) - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) && (353)\\
\leq & n(\lambda_{pp}+\lambda_{np})\log P + I(W_1; Y_{pn}^n, Y_{nn}^n|\Omega) - h(Y_{np}^n|Y_{nn}^n, Y_{pn}^n, W_1, \Omega) + no(n) + no(\log P) && (354)\\
\leq & n(\lambda_{pp}+\lambda_{np})\log P + I(W_1; Y_{pn}^n, Y_{nn}^n, Z_{pn}^n, \tilde{Z}_{pn}^n, Z_{np}^n, W_2|\Omega)\\
& - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) + no(\log P) && (355)\\
= & n(\lambda_{pp}+\lambda_{np})\log P + I(W_1; Y_{pn}^n, \tilde{Z}_{pn}^n|Y_{nn}^n, Z_{pn}^n, Z_{np}^n, W_2, \Omega)\\
& - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) + no(\log P) && (356)\\
= & n(\lambda_{pp}+\lambda_{np})\log P + h(Y_{pn}^n, \tilde{Z}_{pn}^n|Y_{nn}^n, Z_{pn}^n, Z_{np}^n, W_2, \Omega)\\
& - h(Y_{pn}^n, \tilde{Z}_{pn}^n|Z_{pn}^n, Y_{nn}^n, Z_{np}^n, W_1, W_2, \Omega) - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) + no(\log P) && (357)\\
\leq & n(\lambda_{pp}+\lambda_{np})\log P + h(Y_{pn}^n, \tilde{Z}_{pn}^n|Z_{pn}^n, Z_{np}^n, Y_{nn}^n, W_2, \Omega)\\
& - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) + no(\log P) && (358)\\
= & n(\lambda_{pp}+\lambda_{np})\log P + h(\tilde{Z}_{pn}^n|Z_{pn}^n, Z_{np}^n, Y_{nn}^n, W_2, \Omega) + h(Y_{pn}^n|Z_{pn}^n, \tilde{Z}_{pn}^n, Z_{np}^n, Y_{nn}^n, W_2, \Omega)\\
& - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega) + no(n) + no(\log P) && (359)\\
\leq & n(\lambda_{pp}+\lambda_{np})\log P + h(\tilde{Z}_{pn}^n|Z_{np}^n, Y_{nn}^n, W_2, \Omega) - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega)\\
& + no(n) + no(\log P) && (360)\\
= & n(\lambda_{pp}+\lambda_{np})\log P + h(Z_{pn}^n|Z_{np}^n, Y_{nn}^n, W_2, \Omega) - h(Y_{np}^n|Y_{pn}^n, Y_{nn}^n, W_1, \Omega)\\
& + no(n) + no(\log P), && (361)
\end{aligned}$$

where (356) follows since,

$$\begin{aligned}
I(W_1; Z_{pn}^n, Z_{np}^n, Y_{nn}^n, W_2|\Omega) \leq & I(W_1; Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n, W_2|\Omega) && (362)\\
= & I(W_1, Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n|\Omega) + I(W_1; W_2|Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n, \Omega) && (363)\\
= & no(\log P) + I(W_1; W_2|Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n, \Omega) && (364)\\
\leq & no(\log P) + H(W_2|Z_{pp}^n, Z_{pn}^n, Z_{np}^n, Y_{nn}^n, \Omega) && (365)
\end{aligned}$$

$$\leq no(\log P) + no(n), \tag{366}$$

where, (364) and (366) follow from the secrecy and decodability requirements, respectively. In addition, (358) follows since $h(Y_{pn}^n, \tilde{Z}_{pn}^n | Z_{pn}^n, Z_{np}^n, Y_{nn}^n, W_1, W_2, \Omega) \geq o(\log P)$, (360) follows since given Z_{pn}^n and $\tilde{Z}_{pn}^n$, one can reconstruct $\mathbf{X}_{pn}^n$ and hence Y_{pn}^n to within noise distortion, and (361) follows due to the statistical equivalence of receivers 2 and $\tilde{2}$ in the state PN.

Similarly, by symmetry, we have,

$$\begin{aligned} nR_2 \leq & n(\lambda_{pp} + \lambda_{pn}) \log P + h(Y_{np}^n | Y_{pn}^n, Y_{nn}^n, W_1, \Omega) \\ & - h(Z_{pn}^n | Z_{np}^n, Y_{nn}^n, W_2, \Omega) + no(n) + no(\log P). \end{aligned} \tag{367}$$

Adding (361) and (367), we have,

$$n(R_1 + R_2) \leq n(2\lambda_{pp} + \lambda_{pn} + \lambda_{np}) \log P + 2no(n) + o(\log P). \tag{368}$$

First dividing by $n \log(P)$ and letting $n \to \infty$, and then letting $P \to \infty$, we obtain,

$$d_1 + d_2 \leq 2\lambda_{pp} + \lambda_{pn} + \lambda_{np}. \tag{369}$$

This completes the proof of Lemma 4.

C.4 Proof of Lemma 5

We want to show that for the two-user MISO BC with only four of the nine states: PP, PD, DP and DD occurring for λ_{pp}, λ_{pd}, λ_{dp} and λ_{dd} fractions of the time, with $\lambda_{pd} = \lambda_{dp}$ and $\lambda_{pp} + \lambda_{pd} + \lambda_{dp} + \lambda_{dd} = 1$,

$$3d_1 + d_2 \leq 2 + 2\lambda_{pp} + 2\lambda_{pd} \tag{370}$$

$$d_1 + 3d_2 \leq 2 + 2\lambda_{pp} + 2\lambda_{pd}. \tag{371}$$

To do so, for each of the two receivers, we introduce another statistically equivalent receiver. At receiver 1, we introduce a virtual receiver $\tilde{1}$, with channel output denoted by $\tilde{Y}$, while the channel output at the virtual receiver $\tilde{2}$ at receiver 2 is denoted by $\tilde{Z}$. Since the capacity depends on the marginals, without loss of generality, we can assume that the channels in the state NN are the same for all receivers. The outputs at each of the receivers can be written as

$$Y^n = (Y_{pp}^n, Y_{pd}^n, Y_{dp}^n, Y_{nn}^n) \tag{372}$$

$$Z^n = (Z_{pp}^n, Z_{pd}^n, Z_{dp}^n, Y_{nn}^n) \tag{373}$$

$$\tilde{Y}^n = (Y_{pp}^n, Y_{pd}^n, \tilde{Y}_{dp}^n, Y_{nn}^n) \tag{374}$$
$$\tilde{Z}^n = (Z_{pp}^n, \tilde{Z}_{pd}^n, Z_{dp}^n, Y_{nn}^n), \tag{375}$$

where

$$\tilde{Y}_{dp}(t) = \tilde{\mathbf{H}}_{1,dp}(t)\mathbf{X}_{dp}(t) + \tilde{N}_{1,dp}(t) \tag{376}$$
$$\tilde{Z}_{pd}(t) = \tilde{\mathbf{H}}_{2,pd}(t)\mathbf{X}_{pd}(t) + \tilde{N}_{2,pd}(t), \tag{377}$$

such that $\tilde{\mathbf{H}}_{1,dp}$, $\tilde{\mathbf{H}}_{2,pd}$ are i.i.d. with the same distribution as $\mathbf{H}_{1,dp}$, $\mathbf{H}_{2,pd}$, respectively, and $\tilde{N}_{1,dp}$, $\tilde{N}_{2,pd}$ are i.i.d. with same distribution as $N_{1,dp}$, $N_{2,pd}$. We consider a special case with only four states PP, PD, DP and DD. Aided by Lemma 1, we proceed to prove Lemma 5, as follows:

$$\begin{aligned}
nR_1 \leq & I(W_1; Y^n|\Omega) + no(n) && (378)\\
\leq & I(W_1; Y^n|\Omega) - I(W_1; Z_{dp}^n Z_{dd}^n|\Omega) + no(\log P) + no(n) && (379)\\
\leq & h(Y^n|\Omega) - \frac{1}{2}h(Z_{dp}^n, Z_{dd}^n|W_1, \Omega) - h(Z_{dp}^n, Z_{dd}^n|\Omega) + h(Z_{dp}^n, Z_{dd}^n|W_1, \Omega) \\
& + no(\log P) + no(n) && (380)\\
= & h(Y^n|\Omega) + \frac{1}{2}h(Z_{dp}^n, Z_{dd}^n|W_1, \Omega) - h(Z_{dp}^n, Z_{dd}^n|\Omega) + no(\log P) + no(n) && (381)\\
\leq & h(Y^n|\Omega) + \frac{1}{2}h(Z_{dp}^n, Z_{dd}^n|\Omega) - h(Z_{dp}^n, Z_{dd}^n|\Omega) + no(\log P) + no(n) && (382)\\
= & h(Y^n|\Omega) - \frac{1}{2}h(Z_{dp}^n, Z_{dd}^n|\Omega) + no(\log P) + no(n) && (383)\\
\leq & n\log P - \frac{1}{2}h(Z_{dp}^n, Z_{dd}^n|\Omega) + no(\log P) + no(n), && (384)
\end{aligned}$$

where (379) follows from the security constraints, (380) follows from a conditioned version of Lemma 1 (conditioned on W_1), and (382) follows, since conditioning reduces differential entropy.

We also have the following bounds for user 1:

$$\begin{aligned}
nR_1 \leq & I(W_1; Y^n|W_2, \Omega) + no(n) && (385)\\
\leq & I(W_1; Y^n, Z^n|W_2, \Omega) + no(n) && (386)\\
= & I(W_1; Y^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (387)\\
\leq & h(Y^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (388)\\
= & h(Y_{pd}^n, Y_{dp}^n, Y_{dd}^n|Z^n, W_2, \Omega) + h(Y_{pp}^n|Y_{pd}^n, Y_{dp}^n, Y_{dd}^n, Z^n, W_2, \Omega) + no(\log P) + no(n) && (389)\\
\leq & n\lambda_{pp}\log P + h(Y_{dp}^n|Z^n, W_2, \Omega) + h(Y_{pd}^n, Y_{dd}^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (390)\\
\leq & n(\lambda_{pp} + \lambda_{dp})\log P + h(Y_{pd}^n, Y_{dd}^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (391)
\end{aligned}$$

$$\leq n(\lambda_{pp}+\lambda_{dp})\log P + h(Z^n|W_2,\Omega) + no(\log P) + no(n), \tag{392}$$

where (387) follows since,

$$\begin{aligned} I(W_1;Z^n|W_2,\Omega) \leq & I(W_1;Z^n,W_2|\Omega) && (393)\\ = & I(W_1;Z^n|\Omega) + I(W_1;W_2|Z^n,\Omega) && (394)\\ \leq & no(\log P) + H(W_2|Z^n,\Omega) && (395)\\ \leq & no(\log P) + no(n), && (396)\end{aligned}$$

using the security and reliability constraints. In addition, (392) follows from the conditional version of Lemma 1 (conditioned on W_2).

For receiver 2, we have

$$\begin{aligned} nR_2 \leq & I(W_2;Z^n|\Omega) + no(n) && (397)\\ = & h(Z^n|\Omega) - h(Z^n|W_2,\Omega) + no(n) && (398)\\ = & h(Z_{pp}^n|Z_{pd}^n,Z_{dp}^n,Z_{dd}^n,\Omega) + h(Z_{pd}^n,Z_{dp}^n,Z_{dd}^n|\Omega) - h(Z^n|W_2,\Omega) + no(n) && (399)\\ \leq & n\lambda_{pp}\log P + h(Z_{pd}^n|\Omega) + h(Z_{dp}^n,Z_{dd}^n|\Omega) - h(Z^n|W_2,\Omega) + no(n) && (400)\\ \leq & n(\lambda_{pp}+\lambda_{dp})\log P + h(Z_{dp}^n,Z_{dd}^n|\Omega) - h(Z^n|W_2,\Omega) + no(n). && (401)\end{aligned}$$

In summary, from (384), (392) and (401), we have,

$$\begin{aligned} nR_1 \leq & n\log P - \frac{1}{2}h(Z_{dp}^n,Z_{dd}^n|\Omega) + no(\log P) + no(n), && (402)\\ nR_1 \leq & n(\lambda_{pp}+\lambda_{dp})\log P + h(Z^n|W_2,\Omega) + no(\log P) + no(n), && (403)\\ nR_2 \leq & n(\lambda_{pp}+\lambda_{dp})\log P + h(Z_{dp}^n,Z_{dd}^n|\Omega) - h(Z^n|W_2,\Omega) + no(n). && (404)\end{aligned}$$

Eliminating $h(Z_{dp}^n,Z_{dd}^n|\Omega)$ and $h(Z^n|W_2,\Omega)$ from these inequalities and taking the limit $n\to\infty$, we arrive at

$$3R_1 + R_2 \leq (2+2\lambda_{pp}+2\lambda_{dp})\log P + o(\log P). \tag{405}$$

Dividing by $\log P$ and taking the limit $P\to\infty$, we get the required result

$$3d_1 + d_2 \leq 2 + 2\lambda_{pp} + 2\lambda_{dp}. \tag{406}$$

Appendix D Proof of the s.d.o.f. Region for PD State

In this section, we present the proof for the s.d.o.f. region of the fixed PD state (perfect CSIT from user 1 and delayed CSIT from user 2). The s.d.o.f. region in this case is given by all

non-negative pairs (d_1, d_2) satisfying,

$$d_1 + d_2 \leq 1. \tag{407}$$

To prove this claim, we first provide a proof of the converse and then two achievable schemes that are sufficient to achieve the full region.

D.1 Converse

To this end, we create a virtual receiver with output $\tilde{Z}^n$ with a channel that is statistically equivalent to user 2. The channel output $\tilde{Z}$ is given by

$$\tilde{Z}(t) = \tilde{\mathbf{H}}_2(t)\mathbf{X}(t) + \tilde{N}_2(t), \tag{408}$$

where $\tilde{\mathbf{H}}_2$ and $\tilde{N}_2$ are i.i.d. as $\mathbf{H}_2$ and N_2, respectively. Then, the local statistical equivalence property implies that

$$h(Z(t)|Z^{t-1}, W_2, \Omega) = h(\tilde{Z}(t)|Z^{t-1}, W_2, \Omega), \tag{409}$$

where Ω is the set of all channel coefficients upto and including time n. Let us now bound the rate of user 1:

$$\begin{aligned}
nR_1 \leq & I(W_1; Y^n|W_2, \Omega) + no(n) && (410)\\
\leq & I(W_1; Y^n, Z^n|W_2, \Omega) + no(n) && (411)\\
= & I(W_1; Y^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (412)\\
\leq & I(W_1; Y^n, \tilde{Z}^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (413)\\
= & h(Y^n, \tilde{Z}^n|Z^n, W_2, \Omega) - h(Y^n, \tilde{Z}^n|Z^n, W_1, W_2, \Omega) + no(\log P) + no(n) && (414)\\
\leq & h(Y^n, \tilde{Z}^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (415)\\
= & h(\tilde{Z}^n|Z^n, W_2, \Omega) + h(Y^n|Z^n, \tilde{Z}^n, W_2, \Omega) + no(\log P) + no(n) && (416)\\
\leq & h(\tilde{Z}^n|Z^n, W_2, \Omega) + no(\log P) + no(n) && (417)\\
= & \sum_{t=1}^{n} h(\tilde{Z}(t)|\tilde{Z}^{t-1}, Z^n, W_2, \Omega) + no(\log P) + no(n) && (418)\\
\leq & \sum_{t=1}^{n} h(\tilde{Z}(t)|Z^{t-1}, W_2, \Omega) + no(\log P) + no(n) && (419)\\
= & \sum_{t=1}^{n} h(Z(t)|Z^{t-1}, W_2, \Omega) + no(\log P) + no(n) && (420)\\
= & h(Z^n|W_2, \Omega) + no(\log P) + no(n), && (421)
\end{aligned}$$

where (412) follows since $I(W_1; Z^n|W_2, \Omega) \leq no(\log P)$ from (393), (417) follows due to the fact that given Z^n and $\tilde{Z}^n$, it is possible to reconstruct $\mathbf{X}^n$ and hence Y^n to within noise distortion, and (420) follows from (409).

For the second user, we have,

$$\begin{aligned} nR_2 \leq & I(W_2; Z^n|\Omega) + no(n) && (422)\\ = & h(Z^n|\Omega) - h(Z^n|W_2, \Omega) + no(n) && (423)\\ \leq & n \log P - h(Z^n|W_2, \Omega) + no(n). && (424) \end{aligned}$$

Adding (421) and (424), we have,

$$n(R_1 + R_2) \leq n \log P + no(\log P) + no(n). \tag{425}$$

Dividing by n and letting $n \to \infty$,

$$R_1 + R_2 \leq \log P + o(\log P). \tag{426}$$

Now dividing by $\log P$ and letting $P \to \infty$,

$$d_1 + d_2 \leq 1. \tag{427}$$

This completes the proof of the converse for the case of PD state alone.

D.2 Achievable Schemes

Note that it is sufficient to achieve only two points: a) $(d_1, d_2) = (1, 0)$ and b) $(d_1, d_2) = (0, 1)$. The achievability of these corner points follow in straightforward manner from existing arguments as follows: sending message to user 1 by superimposing it with artificial noise in a direction orthogonal to user 1's channel to achieve the pair $(1, 0)$; and sending the message to user 2 in a direction orthogonal to user 1's channel to achieve the pair $(0, 1)$. This completes the proof of the achievability of the region in (407).

Appendix E Proof of the s.d.o.f. Region for DN State

For the MISO BCCM with the fixed state DN (delayed CSIT from the first user and no CSIT from the second user), the s.d.o.f. region is given by the set of all non-negative pairs (d_1, d_2) satisfying,

$$d_1 + d_2 \leq \frac{1}{2}. \tag{428}$$

To prove this claim, we first provide a proof of the converse and then two achievable schemes that are sufficient to achieve the full region.

E.1 Converse

We first create a virtual receiver with output $\tilde{Y}^n$ with a statistically equivalent channel as user 1. The channel output $\tilde{Y}(t)$ is given by

$$\tilde{Y}(t) = \tilde{\mathbf{H}}_1(t)\mathbf{X}(t) + \tilde{N}_1(t), \tag{429}$$

where $\tilde{\mathbf{H}}_1$ and $\tilde{N}_1$ are i.i.d. as $\mathbf{H}_1$ and N_1, respectively. Then, the local statistical equivalence property implies that

$$h(Y(t)|Y^{t-1}, W_1, \Omega) = h(\tilde{Y}(t)|Y^{t-1}, W_1, \Omega), \tag{430}$$

where Ω is the set of all channel coefficients upto and including time n. Similar to the proof of Lemma 1, Appendix B, it can be readily shown that,

$$2h(Y^n|W_1, \Omega) \geq h(Z^n|W_1, \Omega) + o(\log P). \tag{431}$$

Then, for the first user, we have,

$$\begin{aligned}
nR_1 \leq & I(W_1; Y^n|\Omega) - I(W_1; Z^n|\Omega) + no(n) + no(\log P) && (432)\\
= & h(Y^n|\Omega) - h(Y^n|W_1, \Omega) - h(Z^n|\Omega) + h(Z^n|W_1, \Omega) && (433)\\
\leq & h(Y^n|\Omega) - \frac{1}{2}h(Z^n|W_1, \Omega) - h(Z^n|\Omega) + h(Z^n|W_1, \Omega) && (434)\\
= & h(Y^n|\Omega) + \frac{1}{2}h(Z^n|W_1, \Omega) - h(Z^n|\Omega) && (435)\\
\leq & h(Y^n|\Omega) + \frac{1}{2}h(Z^n|\Omega) - h(Z^n|\Omega) && (436)\\
= & h(Y^n|\Omega) - \frac{1}{2}h(Z^n|\Omega), && (437)
\end{aligned}$$

where (434) follows from (431). For the second user,

$$\begin{aligned}
nR_2 \leq & I(W_2; Z^n|\Omega) - I(W_2; Y^n|\Omega) + no(n) + no(\log P) && (438)\\
= & h(Z^n|\Omega) - h(Y^n|\Omega) + (h(Y^n|W_2, \Omega) - h(Z^n|W_2, \Omega)) + no(n) + no(\log P). && (439)
\end{aligned}$$

Adding (437) and (439), we obtain,

$$n(R_1 + R_2) \leq \frac{1}{2}h(Z^n|\Omega) + (h(Y^n|W_2, \Omega) - h(Z^n|W_2, \Omega)) + no(n) + no(\log P) \tag{440}$$

$$\leq \frac{n}{2}\log P + (h(Y^n|W_2,\Omega) - h(Z^n|W_2,\Omega)) + no(n) + no(\log P). \tag{441}$$

Thus, in order to obtain $d_1+d_2 \leq 1/2$, it suffices to show that $(h(Y^n|W_2,\Omega) - h(Z^n|W_2,\Omega)) \leq no(\log P)$, where the transmitter has delayed CSIT from user 1 and no CSIT from user 2. To this end, we invoke a recent result in [47, (39)-(66)], which showed that the maximum of $h(Y^n|W_2,\Omega) - h(Z^n|W_2,\Omega)$ is less than $no(\log P)$, under the assumption of perfect CSIT from user 1 and no CSIT from user 2. Hence, the same upper bound on the maximum value also holds under a weaker assumption of delayed CSIT from user 1. Thus, using the fact that

$$(h(Y^n|W_2,\Omega) - h(Z^n|W_2,\Omega)) \leq no(\log P), \tag{442}$$

and substituting in (441), we have,

$$n(R_1+R_2) \leq \frac{n}{2}\log P + no(n) + no(\log P). \tag{443}$$

Dividing by n and letting $n \to \infty$, we get,

$$R_1 + R_2 \leq \frac{1}{2}\log P + o(\log P). \tag{444}$$

Dividing by $\log P$ and letting $P \to \infty$ yields

$$d_1 + d_2 \leq \frac{1}{2}. \tag{445}$$

This completes the proof of the converse.

E.2 Achievable Schemes

To prove the achievability of the s.d.o.f. region in (428), it suffices to consider only the two points: a) $(d_1,d_2) = \left(\frac{1}{2},0\right)$ and b) $(d_1,d_2) = \left(0,\frac{1}{2}\right)$. Every other point in the region can be obtained by time-sharing. A scheme for achieving $(d_1,d_2) = \left(\frac{1}{2},0\right)$ was presented in [43]. We include it here for completeness.

E.2.1 Scheme Achieving $(d_1,d_2) = \left(\frac{1}{2},0\right)$:

We wish to send 1 symbol u securely to the first user in 2 time slots. This can be done as follows:

1) At time $t = 1$: The transmitter does not have any channel knowledge. It sends:

$$\mathbf{X}(1) = [q_1 \quad q_2]^T, \tag{446}$$

where q_1 and q_2 denote independent artificial noise symbols distributed as $\mathcal{CN}(0,P)$. Both receivers receive linear combinations of the two symbols q_1 and q_2. The receivers' outputs are:

$$Y(1) = h_{11}(1)q_1 + h_{12}(1)q_2 \triangleq L_1(q_1,q_2) \tag{447}$$

$$Z(1) = h_{21}(1)q_1 + h_{22}(1)q_2. \tag{448}$$

Due to delayed CSIT from receiver 1, the transmitter can reconstruct $L_1(q_1,q_2)$ in the next time slot and use it for transmission.

2) At time $t=2$: The transmitter sends:

$$\mathbf{X}(2) = [u \quad L_1(q_1,q_2)]^T. \tag{449}$$

The received signals are:

$$Y(2) = h_{11}(2)u + h_{12}(2)L_1(q_1,q_2) \tag{450}$$

$$Z(2) = h_{21}(2)u + h_{22}(2)L_1(q_1,q_2). \tag{451}$$

Since the receivers have full channel knowledge, receiver 1 can recover u by eliminating $L_1(q_1,q_2)$ from Y(1) and Y(2). On the other hand, the information leakage to the second user is given by,

$$I(u;Z(1),Z(2)|\Omega) = h(Z(1),Z(2)|\Omega) - h(Z(1),Z(2)|u,\Omega) \tag{452}$$

$$\leq 2\log P - h(h_{21}(1)q_1 + h_{22}(1)q_2, h_{11}(1)q_1 + h_{12}(1)q_2|\Omega) \tag{453}$$

$$= 2\log P - 2\log P + o(\log P) \tag{454}$$

$$= o(\log P). \tag{455}$$

E.2.2 Scheme Achieving $(d_1,d_2) = \left(0,\frac{1}{2}\right)$:

In this scheme, we wish to send 1 symbol u securely to the second user in 2 time slots. This can be done as follows:

1) At time $t=1$: The transmitter does not have any channel knowledge. It sends:

$$\mathbf{X}(1) = [u \quad q_1]^T, \tag{456}$$

where q denotes an independent artificial noise symbol distributed as $\mathcal{CN}(0,P)$. Both receivers receive linear combinations of the two symbols u and q. The receivers' outputs are:

$$Y(1) = h_{11}(1)u + h_{12}(1)q \triangleq L(u,q) \tag{457}$$

$$Z(1) = h_{21}(1)u + h_{22}(1)q \triangleq G(u,q). \tag{458}$$

Due to delayed CSIT from receiver 1, the transmitter can reconstruct $L(u,q)$ in the next times lot and use it for transmission.

2) At time $t = 2$: The transmitter sends:

$$\mathbf{X}(2) = \begin{bmatrix} L(u,q) & 0 \end{bmatrix}^T. \tag{459}$$

The received signals are:

$$Y(2) = h_{11}(2)L(u,q) \tag{460}$$
$$Z(2) = h_{21}(2)L(u,q). \tag{461}$$

Since the receivers have full channel knowledge, receiver 2 can recover u by eliminating q from $L(u,q)$ and $G(u,q)$. On the other hand, the information leakage to the first user is given by,

$$I(u; Y(1), Y(2)|\Omega) = I(u; L(u,q)|\Omega) \tag{462}$$
$$= h(L(u,q)|\Omega) - h(L(u,q)|u,\Omega) \tag{463}$$
$$\leq \log P - \log P + o(\log P) \tag{464}$$
$$= o(\log P). \tag{465}$$

This completes the proof of achievability.

References

[1] Y. Liang, H. V. Poor, and S. Shamai. Secure communication over fading channels. *IEEE Transactions on Information Theory*, 54(6):2470–2492, Jun. 2008.

[2] Z. Li, R. D. Yates, and W. Trappe. Secrecy capacity of independent parallel channels. In R. Liu and W. Trappe, editors, *Securing Wireless Communications at the Physical Layer*, pages 1–18. Springer US, 2010.

[3] P. K. Gopala, L. Lai, and H. El Gamal. On the secrecy capacity of fading channels. *IEEE Transactions on Information Theory*, 54(10):4687–4698, Oct. 2008.

[4] P. Mukherjee and S. Ulukus. Fading wiretap channel with no CSI anywhere. In Proc. *IEEE International Symposium on Information Theory*, pages 1347–1351, Jul. 2013.

[5] S. Shafiee, N. Liu, and S. Ulukus. Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel. *IEEE Transactions on Information Theory*, 55(9):4033–4039, Sept. 2009.

[6] A. Khisti and G. W. Wornell. Secure transmission with multiple antennas - Part II: The MIMOME wiretap channel. *IEEE Transactions on Information Theory*, 56(11):5515–5532, Nov. 2010.

[7] F. Oggier and B. Hassibi. The secrecy capacity of the MIMO wiretap channel. *IEEE Transactions on Information Theory*, 57(8):4961–4972, Aug. 2011.

[8] T. Liu and S. Shamai. A note on the secrecy capacity of the multiple-antenna wiretap channel. *IEEE Transactions on Information Theory*, 55(6):2547–2553, Jun. 2009.

[9] E. Tekin and A. Yener. The Gaussian multiple access wire-tap channel. *IEEE Transactions on Information Theory*, 54(12):5747–5755, Dec. 2008.

[10] E. Tekin and A. Yener. The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming. *IEEE Transactions on Information Theory*, 54(6):2735–2751, June 2008.

[11] E. Ekrem and S. Ulukus. On the secrecy of multiple access wiretap channel. In Proc. *46th annual Allerton Conference on Communication, Control, and Computing*, Sept. 2008.

[12] R. Bassily and S. Ulukus. Ergodic secret alignment. *IEEE Transactions on Information Theory*, 58(3):1594–1611, March 2012.

[13] J. Xie and S. Ulukus. Secure degrees of freedom of the Gaussian multiple access wiretap channel. In Proc. *IEEE International Symposium on Information Theory*, Jul. 2013.

[14] E. Ekrem and S. Ulukus. The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel. *IEEE Transactions on Information Theory*, 57(4):2083–2114, Apr. 2011.

[15] E. Ekrem and S. Ulukus. Secrecy capacity of a class of broadcast channels with an eavesdropper. *EURASIP Journal on Wireless Communications and Networking*, March 2009.

[16] G. Bagherikaram, A. S. Motahari, and A. K. Khandani. Secure broadcasting: The secrecy rate region. In Proc. *46th annual Allerton Conference on Communication, Control, and Computing*, pages 834–841, Sept 2008.

[17] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates. Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions. *IEEE Transactions on Information Theory*, 54(6):2493–2507, Jun. 2008.

[18] R. Liu and H. V. Poor. Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages. *IEEE Transactions on Information Theory*, 55(3):1235–1249, Mar. 2009.

[19] R. Liu, T. Liu, H. V. Poor, and S. Shamai. Multiple-input multiple-output Gaussian broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 56(9):4215–4227, Sept. 2010.

[20] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor. The Gaussian wiretap channel with a helping interferer. In Proc. *IEEE International Symposium on Information Theory*, pages 389–393, July 2008.

[21] J. Xie and S. Ulukus. Secure degrees of freedom of the Gaussian wiretap channel with helpers. In Proc. *52nd annual Allerton Conference on Communication, Control, and Computing*, Oct. 2012.

[22] O. O. Koyluoglu, H. El Gamal, L. Lai, and H. V. Poor. Interference alignment for secrecy. *IEEE Transactions on Information Theory*, 57(6):3323–3332, June 2011.

[23] O. O. Koyluoglu and H. El Gamal. Cooperative encoding for secrecy in interference channels. *IEEE Transactions on Information Theory*, 57(9):5682–5694, Sept 2011.

[24] X. He and A. Yener. The Gaussian many-to-one interference channel with confidential messages. *IEEE Transactions on Information Theory*, 57(5):2730–2745, May 2011.

[25] J. Xie and S. Ulukus. Unified secure DoF analysis of K-user Gaussian interference channels. In Proc. *IEEE International Symposium on Information Theory*, pages 1107–1111, July 2013.

[26] T. Gou and S. A. Jafar. On the secure degrees of freedom of wireless X networks. In Proc. *46th annual Allerton Conference on Communication, Control, and Computing*, pages 826–833, Sept 2008.

[27] Z. Wang, M. Xiao, M. Skoglund, and H. V. Poor. Secrecy degrees of freedom of wireless X networks using artificial noise alignment. Available at arXiv:1410.5009v2.

[28] L. Lai and H. El Gamal. The relay-eavesdropper channel: Cooperation for secrecy. *IEEE Transactions on Information Theory*, 54(9):4005–4019, Sept 2008.

[29] M. Yuksel and E. Erkip. The relay channel with a wire-tapper. In Proc. *41st Annual Conference on Information Sciences and Systems*, pages 13–18, March 2007.

[30] M. Bloch and A. Thangaraj. Confidential messages to a cooperative relay. In Proc. *IEEE Information Theory Workshop*, pages 154–158, May 2008.

[31] X. He and A. Yener. Cooperation with an untrusted relay: A secrecy perspective. *IEEE Transactions on Information Theory*, 56(8):3807–3827, Aug 2010.

[32] E. Ekrem and S. Ulukus. Secrecy in cooperative relay broadcast channels. *IEEE Transactions on Information Theory*, 57(1):137–155, Jan 2011.

[33] C. S. Vaze and M. K. Varanasi. The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT. *IEEE Transactions on Information Theory*, 58(8):5354–5374, Aug. 2012.

[34] M. A. Maddah-Ali and D. Tse. Completely stale transmitter channel state information is still useful. *IEEE Transactions on Information Theory*, 58(7):4418–4431, Jul. 2012.

[35] C. S. Vaze and M. K. Varanasi. The degrees of freedom regions of two-user and certain three-user MIMO broadcast channels with delayed CSIT. *arXiv:1101.0306v2[cs.IT]*, 2011.

[36] R. Tandon, S. Mohajer, H. V. Poor, and S. Shamai. Degrees of freedom region of the MIMO interference channel with output feedback and delayed CSIT. *IEEE Transactions on Information Theory*, 59(3):1444–1457, Mar. 2013.

[37] S. A. Jafar and S. Shamai. Degrees of freedom region of the MIMO X channel. *IEEE Transactions on Information Theory*, 54(1):151–170, Jan 2008.

[38] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani. Communication over MIMO X-channels: Interference alignment, decomposition, and performance analysis. *IEEE Transactions on Information Theory*, 54(8):3457–3470, Aug 2008.

[39] S. Lashgari, S. Avestimehr, and C. Suh. Linear degrees of freedom of the X-channel with delayed CSIT. *IEEE Transactions on Information Theory*, 60(4):2180–2189, Apr. 2014.

[40] D. T. H. Kao and S. Avestimehr. Linear degrees of freedom of the MIMO X-channel with delayed CSIT. In Proc. *IEEE International Symposium on Information Theory*, pages 366–370, Jun. 2014.

[41] A. Ghasemi, A. S. Motahari, and A. K. Khandani. On the degrees of freedom of X-channel with delayed CSIT. In Proc. *IEEE International Symposium on Information Theory*, pages 767–770, Jul. 2011.

[42] R. Tandon, S. Mohajer, H. V. Poor, and S. Shamai. On X-channels with feedback and delayed CSI. In Proc. *IEEE International Symposium on Information Theory*, pages 1877–1881, Jul. 2012.

[43] S. Yang, M. Kobayashi, P. Piantanida, and S. Shamai. Secrecy degrees of freedom of MIMO broadcast channels with delayed CSIT. *IEEE Transactions on Information Theory*, 59(9):5244–5256, Sept. 2013.

[44] A. Zaidi, Z. H. Awan, S. Shamai, and L. Vandendorpe. Secure degrees of freedom of MIMO X-channels with output feedback and delayed CSIT. *IEEE Transactions on Information Forensics and Security*, 8(11):1760–1774, Nov 2013.

[45] H. Maleki, S. A. Jafar, and S. Shamai. Retrospective interference alignment over interference networks. *IEEE Journal of Selected Topics in Signal Processing*, 6(3):228–240, Jun. 2012.

[46] R. Tandon, M. A. Maddah-Ali, A. Tulino, H. V. Poor, and S. Shamai. On fading broadcast channels with partial channel state information at the transmitter. In Proc. *IEEE International Symposium on Wireless Communication Systems*, Aug. 2012.

[47] A. G. Davoodi and S. A. Jafar. Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT. Available at arXiv:1403.1541v1.

[48] S. Amuru, R. Tandon, and S. Shamai. On the degrees-of-freedom of the 3-user MISO broadcast channel with hybrid CSIT. In Proc. *IEEE International Symposium on Information Theory*, pages 2137–2141, Jun. 2014.

[49] K. Mohanty and M. K. Varanasi. On the dof region of the K-user MISO broadcast channel with hybrid CSIT. Available at arXiv:1311.6647v1.

[50] R. Tandon, S. A. Jafar, S. Shamai, and H. V. Poor. On the synergistic benefits of alternating CSIT for the MISO broadcast channel. *IEEE Transactions on Information Theory*, 59(7):4106–4128, Jul. 2013.

[51] P. Mukherjee, R. Tandon, and S. Ulukus. MISO broadcast channels with confidential messages and alternating CSIT. In Proc. *IEEE International Symposium on Information Theory*, pages 216–220, June 2014.

[52] Z. H. Awan, A. Zaidi, and A. Sezgin. Achievable secure degrees of freedom of MISO broadcast channel with alternating CSIT. In Proc. *IEEE International Symposium on Information Theory*, pages 31–35, June 2014.

[53] A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, 54(8):1355–1387, Oct. 1975.

[54] I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978.